£21
6.8.70

THE
BRITISH JOURNAL
OF
PSYCHOLOGY

4782 69

GENERAL SECTION

EDITED BY
D. W. HARDING
WITH THE ASSISTANCE OF
C. BURT

AIDED IN THE SELECTION OF PAPERS BY
F. C. BARTLETT
E. A. BOTT
R. S. CREED
G. C. DREW
L. S. HEARNSHAW
M. G. KENDALL
T. H. PEAR
G. RYLE
W. J. H. SPROTT
R. H. THOULESS
O. L. ZANGWILL

ASSISTANT EDITOR: J. W. REEVES

VOLUME XLV, 1954

BUREAU OF EDNL. & PSYL. RESEARCH.
LIBRARY.

CAMBRIDGE
AT THE UNIVERSITY PRESS
1954

Printed in Great Britain at the University Press, Cambridge
(Brooke Crutchley, University Printer)
and published by the Cambridge University Press
London: Bentley House, N.W. 1
American Branch: New York
Agents for Canada, India, and Pakistan: Macmillan

Bureau Edni. Psy. Research
DAVID HA ING COLLEGE
Dated 6.8.70
Acc. No. J 21

CONTENTS OF VOL. XLV

PART 1. FEBRUARY 1954

PAGE

The psychological study of conceptual thinking. By L. S. HEARNSHAW . 1

Changes of performance with age on similar tasks of varying complexity. By HILARY M. CLAY 7

Memory mechanisms and the theory of schemata. By R. C. OLDFIELD . 14

The continuous regulation of skilled response by kinaesthetic feed back. By C. B. GIBBS 24

The adequacy of everyday quantitative expressions as measurements of qualities. By D. SHEPPARD 40

Verbal extrapolation and interpolation. By E. C. POULTON 51

A preliminary investigation of the effects of continuous stimulation on the perception of 'apparent motion'. By M. B. SHAPIRO 58

Correspondence 69

Publications recently received 71

PART 2. MAY 1954

Critical notice and appreciation of the work of the late Professor Clark L. Hull. By F. V. SMITH 77

The effect of the vocational aims of industrial apprentices upon their attitude to education and commissioning. By SYLVIA ANTHONY 82

The effect of reference group functions on social status ratings. By GURTH HIGGIN 88

On the variability of the speed of talking and on its relation to the length of utterances in conversations. By FRIEDA GOLDMAN-EISLER 94

An experimental study of comparative judgements of time. By JOHN COHEN, C. E. M. HANSEL and JOHN SYLVESTER 108

Some problems of anomalous colour vision. By R. W. PICKFORD 115

A note on Eysenck's article 'The effects of psychotherapy: an evaluation'. By LESTER LUBORSKY 129

A reply to Luborsky's note. By H. J. EYSENCK 132

Correspondence 134

Publications recently received 138

PART 3. AUGUST 1954

PAGE

Expectancy in apparent visual movement. By E. E. JONES and J. S. BRUNER 157

The change of movement timing with age. By W. T. SINGLETON . . . 166

Cognitive changes in the feebleminded. By A. D. B. CLARKE and A. M. CLARKE 173

A theory of human curiosity. By D. E. BERLYNE 180

A note on Ashanti names and their relationship to personality. By G. JAHODA 192

Evidence of improvement in poetic judgement. By J. N. BRITTON . . 196

A sex difference in preferences for shapes. By W. A. MCELROY . . . 209

A new tachistoscope and cycling timer. By W. C. R. WITHERS . . . 217

Correspondence 220

Critical notice: *Uppsala Symposium on Psychological Factor Analysis, 17–19 March* 1953 223

Publications recently received 225

PART 4. NOVEMBER 1954

Obituary Notice: STANLEY J. F. PHILPOTT, 1888–1952 241

The concept of heed. By U. T. PLACE 243

An experimental study of human curiosity. By D. E. BERLYNE . . . 256

Between wakefulness and sleep: hypnagogic imagery. By PETER MCKELLAR and LORNA SIMPSON 266

Some effects of incentives on the performance of imbeciles. By S. GORDON, N. O'CONNOR and J. TIZARD 277

Children's reactions to bivariant stimuli. By IAN M. L. HUNTER . . . 288

The effect of hunger upon drinking patterns in the rat. By ALLEN D. CALVIN and RICHARD A. BEHAN 294

Publications recently received 299

Other publications received 320

LIST OF AUTHORS

PAGE

Anthony, Sylvia. The effect of the vocational aims of industrial apprentices upon their attitude to education and commissioning 82

Behan, Richard A., *see* Calvin, Allen D.

Berlyne, D. E. A theory of human curiosity 180

Berlyne, D. E. An experimental study of human curiosity 256

Britton, J. N. Evidence of improvement in poetic judgement . . . 196

Bruner, J. S., *see* Jones, E. E.

Calvin, Allen D. and Behan, Richard A. The effect of hunger upon drinking patterns in the rat 294

Clarke, A. D. B. and Clarke, A. M. Cognitive changes in the feebleminded 173

Clarke, A. M., *see* Clarke, A. D. B.

Clay, Hilary M. Changes of performance with age on similar tasks of varying complexity 7

Cohen, John, Hansel, C. E. M. and Sylvester, John. An experimental study of comparative judgements of time 108

Eysenck, H. J. A reply to Luborsky's note 132

Gibbs, C. B. The continuous regulation of skilled response by kinaesthetic feed back 24

Goldman-Eisler, Frieda. On the variability of the speed of talking and on its relation to the length of utterances in conversations 94

Gordon, S., O'Connor, N. and Tizard, J. Some effects of incentives on the performance of imbeciles 277

Hansel, C. E. M., *see* Cohen, John

Hearnshaw, L. S. The psychological study of conceptual thinking . . 1

Higgin, Gurth. The effect of reference group functions on social status ratings 88

Hunter, Ian M. L. Children's reactions to bivariant stimuli . . . 288

Jahoda, G. A note on Ashanti names and their relationship to personality . 192

Jones, E. E. and Bruner, J. S. Expectancy in apparent visual movement . 157

Luborsky, Lester. A note on Eysenck's article 'The effects of psychotherapy: an evaluation' 129

McElroy, W. A. A sex difference in preferences for shapes 209

McKellar, Peter and Simpson, Lorna. Between wakefulness and sleep: hypnagogic imagery 266

O'Connor, N., *see* Gordon, S.

PAGE

OLDFIELD, R. C. Memory mechanisms and the theory of schemata . . . 14

PICKFORD, R. W. Some problems of anomalous colour vision 115

PLACE, U. T. The concept of heed 243

POULTON, E. C. Verbal extrapolation and interpolation 51

SHAPIRO, M. B. A preliminary investigation of the effects of continuous stimulation on the perception of 'apparent motion' 58

SHEPPARD, D. The adequacy of everyday quantitative expressions as measurements of qualities 40

SIMPSON, LORNA, *see* MCKELLAR, PETER

SINGLETON, W. T. The change of movement timing with age 166

SMITH, F. V. Critical notice and appreciation of the work of the late Professor Clark L. Hull 77

SYLVESTER, JOHN, *see* COHEN, JOHN

TIZARD, J., *see* GORDON, S.

WITHERS, W. C. R. A new tachistoscope and cycling timer 217

WOLTERS, A. W. Obituary notice: Stanley J. F. Philpott, 1888–1952 . . 241

VOLUME XLV FEBRUARY 1954 PART 1

4787 6

THE PSYCHOLOGICAL STUDY OF CONCEPTUAL THINKING[1]

BY L. S. HEARNSHAW

Department of Psychology, University of Liverpool

I. *Current interest in conceptual thinking* (pp. 1–3). II. *Problems and methods* (pp. 3–5). *References* (p. 5).

I. CURRENT INTEREST IN CONCEPTUAL THINKING

The psychology of thinking briefly held a prominent place among the interests of psychologists at the beginning of the century, when a group of German investigators, commonly known as the Würzburg school, attempted to analyse human thinking by means of introspection. Prof. Humphrey of Oxford has very ably summarized these investigations in his recent book, *Thinking* (1951), but in spite of his endeavours at rehabilitation the more usual view is that this early work miscarried, primarily because it was attempting an impossible task in trying to unravel complex, rapid and partly subconscious processes by introspective means. At least there can be no doubt that these investigations petered out, and that between the wars experimental investigations into the thought processes of normal adults were few in number. Interest turned to the more experimentally accessible processes of perception and learning, and although intelligence testing thrived prolifically few psychologists, other than Spearman, showed much concern for the nature of the intellect which they were presumably trying to measure.

Since the second World War there has been a growing recognition that the psychology of thinking is not only an important area of psychology but one that can be fruitfully investigated. Hebb (1949), in his already famous book *The Organization of Behaviour*, writes: 'The central problem with which we must find a way to deal...is the problem of thought; some sort of process that is not fully controlled by environmental stimulation and yet cooperates fully with that stimulation. The failure of psychology to handle thought adequately has been the essential weakness of modern psychological theory.' This focusing of interest upon thought processes has been the result of developments in a number of different fields. I propose briefly to review some of the more important of these developments.

In 1934 the Russian psychologist Vigotsky, on the basis of clinical observations and experiments, put forward the theory that a characteristic feature of schizophrenia was a loss of the ability to think in abstract concepts and a regression to a more primitive level which he called 'thinking in complexes'. Vigotsky's work was followed up by Hanfmann & Kasanin (1942) in U.S.A. Using Vigotsky's concept formation test they distinguished two main levels and an intermediate level of performance. At the most primitive level objects are grouped in complexes, and these groupings never acquire generality and never transcend the concrete sphere. 'In forming complexes the subject displays a certain passivity...does not stop to deliberate, to formulate a question or a hypothesis.' At the highest level grouping is carried out on the basis of abstract general concepts, and subjects show spontaneity, and a capacity for forming hypotheses.

---

[1] The first paper in a symposium on conceptual thinking, delivered at the Liverpool meeting of the British Association for the Advancement of Science, September 1953.

Remarkably similar general conclusions were reached by Goldstein & Scheerer (1941) working with brain-injured patients. They distinguished abstract and concrete levels of performance. The clinical psychologist has, therefore, become deeply interested in the psychology of thinking and in methods of diagnosing levels of intellectual performance. Unimpaired thinking in persons of average intelligence is marked among other things by the employment of abstract concepts.

Although we must be careful not to identify precisely the levels of regression described by the psychiatrists with development levels, we can, nevertheless, recognize a certain parallelism between them. Piaget, in particular, has made an intensive study of the developing intelligence of the child, and his main conclusions have been conveniently summarized in *The Psychology of Intelligence* (1950), recently translated into English. In his earlier work, *The Language and Thought of the Child* (1926), Piaget was content to recognize two broad levels of thinking, the egocentric level of the child and the socialized intelligence of the adult. His subsequent investigations have forced him to recognize that 'there is a very long way to go from preverbal intelligence to operational thought before reflective groupings may be established'. The two points of the progression are now termed by Piaget sensori-motor intelligence and conceptual intelligence respectively. His researches have done much to expose the complex foundations upon which conceptual intelligence rests, and to revive the interest of psychologists in the higher thought processes.

A further and perhaps decisive factor in forcing the central processes of thinking upon the attention of psychologists has been the realization by experimentalists working largely with animals that perception and learning are not wholly peripheral and self-sufficient processes, but that even at the animal level they are complicated by generalization and transfer, in other words by rudimentary thought processes. No sharp dividing line can be drawn between generalization and transfer at the sensori-motor level and conceptual thinking. As Piaget points out, 'intelligence is not a completely new power superimposed all of a sudden on completely prepared previous mechanisms, but is only the expression of these same mechanisms when they go beyond present and immediate contact with the world'. Fields (1932) in demonstrating that, after training, rats could eventually discriminate any kind of triangle in any position from other figures went on to suggest that they had attained a concept of triangularity. Lashley's (1934) equivalent stimuli, Pavlov's (1927) stimulus generalization, Harlow's (1949) learning sets lead away from the assumption of a sensory dominance of behaviour and towards the acceptance of autonomous central processes. An important aspect of these central processes is that they are concerned with sorting, classifying, ordering, abstracting—with the operations, therefore, that are characteristic of conceptual thinking. As Hebb has rightly seen, experimental psychology moves towards and requires a theory of thought.

Nor have movements outside psychology itself been without their influence. Craik (1943) was one of the first psychologists to grasp the implication for psychology of the development of mechanical and electrical devices capable of receiving, sorting and using information, and in his book, *The Nature of Explanation*, he treats the brain as in some ways comparable to such devices. The attempt to reproduce mechanically some of the intellectual processes of organisms has necessitated a much more exacting analysis of these processes, and this has again brought to the fore some of the cardinal features of conceptual thinking. For example, when Grey Walter (1953) constructed an electrical mechanism capable of simple learning he found that 'the several learning circuits are really extracting from a selection of events the features that are common to them in time and space—they are in effect recognizing a pattern'. Thus even in quite simple processes there are important analogies to conceptual thinking.

Finally, the philosophers have been approaching what are to them age-old problems with renewed vigour and freshness. Perhaps two recently published works deserve special mention. First, Susanne Langer's (1942) *Philosophy in a New Key*, a brilliant study of symbolism and its functions. 'Symbols', according to her, 'are not proxy for their objects, but are vehicles for the conception of objects...so our interest in the mind has shifted more and more from the acquisition of experience, the domain of sense, to the uses of sense-data, the realm of conception and expression'. The second book is Price's (1953) *Thinking and Experience* which is described by its author as 'an examination of the nature of conceptual cognition, its relation to sense experience on the one hand, and to symbols (verbal and non-verbal) on the other'. Price's analysis of conceptual thinking and his dispositional version of conceptualism are of the highest interest to psychologists. It is indeed in general tenor not unlike accounts given by psychologists themselves, for example by Ribot in his *L'evolution des idées générales* in 1897 or by Wolters in 1933.

Thus by a variety of routes, and for a variety of reasons, the contemporary psychologist has reached the position where he probably would not dissent from William James (1890), who described conception as 'the most important of all features of our mental structure'.

II. PROBLEMS AND METHODS

What are the main problems which confront the psychologist in this field? What methods can he use in his attack on these problems? Is anything of value likely to emerge?

But before we proceed to consider these questions, briefly what is a concept? For the logician abstract general concepts are 'signs pointing to invariant relations, i.e. relations which remain identical despite the variations of the material in which they are embodied. Significant concepts enable us to arrange in order and hold together diverse phenomena because of some real unity of process or relation which constitutes an element of identity between them' (Cohen, 1946). For the psychologist a concept is 'a habit, an organized memory' (Ribot, 1897), or perhaps better as Price (1953) puts it, 'a recognitional capacity'. 'A concept is not before the mind', adds Price, 'as an object of inspection. It is at work *in* the mind...it shows itself not as a detachable item of mental furniture, but rather as a guiding force, determining the direction which the series of presented particulars takes.' Conceptual dispositions may manifest themselves in a great many different ways, and at all levels from the bare capacity to recognize instances in varying contexts to full blown concepts at a high level of abstraction. Taking Humphrey's (1951) definition we may describe conceptual thinking as 'The activity whereby an organism comes to effect a constant modification towards an invariable feature or set of features occurring in a variable context'. But for the psychologist a merely verbal definition is not enough, and, as Heidbreder (1946, 1947–9) has rightly seen, a precise operational definition of conceptual activity which may serve as criterion of its presence is essential.

The psychologist's problems must be clearly distinguished from the traditional problems of the philosopher. The philosopher is mainly concerned with the ontological and logical status of universals. The psychologist is concerned with the *capacity* for conceptual thinking, i.e. for grasping universals. It is only comparatively recently that the psychologist's problem has crystallized. The earlier psychological inquiries were concerned with a phenomenological analysis which was in practice difficult to dissociate from philosophical questions. The question that Aveling asked in 1912, for instance, was 'What is discoverable in consciousness when we think the universal?', and similarly Fisher in 1916 asked 'What are the events which transpire in an observer's consciousness after he has been confronted with a task whose successful performance necessitates the act of generalizing abstraction?' Hull's investigation in 1920 marks an important milestone in the progress of conceptual research. He concentrated on the functional and quantitative aspects of the problem which he pointed out were largely neglected in the earlier introspective and qualitative studies. He employed a small group of abnormal subjects in a subsidiary experiment to investigate individual differences in concept formation and thus initiated research into concept formation as a capacity.

The job of the psychologist, therefore, is to investigate conceptual thinking as a capacity of the organism. The field can be broken down into a number of specific problems. From what beginnings and through what stages, first, does this capacity evolve both in the individual and in the race? This is clearly the problem with which the genetic studies of Piaget are primarily concerned. He has traced the development of conceptual thinking in the child not only in respect of conceptual thinking in general, but in relation to specific key concepts such as those of number, space, time, causality, chance and so on. Anthropological investigation and the study of primitive language provide comparable though less adequate data about the phylogenetic development of conceptual thinking. For the beginnings of conceptual thinking we must have recourse to animal studies for, as Price (1953) points out, there is reason to hold that in a sense conceptual thinking is 'coterminous with conscious life itself'. The phenomena of generalization, transfer and learning sets are met with in the study of animal behaviour, and with Hayek (1952) we may perhaps regard the operation of classifying as a basic feature of nervous mechanisms. Evidently a primary task for the psychologist is to map in the greatest detail the whole evolutionary course of conceptual thinking, and, if the term is to be used as broadly as Price suggests, to distinguish levels of complexity. This provides among other things both the best safeguard against the unduly simplified accounts and explanations of the armchair philosopher, and the necessary material for an analysis of the underlying mechanisms.

Probably no analysis of these mechanisms is going to be satisfactory unless closely tied to neurophysiological facts, and the precise neurophysiological explanation of these complex events still seems some way distant, though as Hebb's speculations suggest the gap is less wide than it was. Phenomenological analysis by introspective means has proved unreliable in practice, and is doubtful in principle. But a behaviouristic analysis of performance under systematically varied conditions such as that carried out by Hull, Heidbreder (1946, 1947–9) and Reed (1946–50) in U.S.A. may indirectly throw light on the mechanisms of conceptual thinking. A number of interesting conclusions have already emerged from this type of research, though at present they have not been integrated into any systematic behaviour theory embracing the whole range of perception, learning and response, nor have they been linked to neurological events. Such an integration is clearly another of the major tasks of the psychologist.

A third group of problems relates to the functional significance of conceptual development. What does this capacity enable its possessor to do? Is it the precondition of flexibility of behaviour, of delayed response, of the development of symbolic systems and language? What, in other words, is its biological value as a mode of adaptation?

Fourthly, what is the place of the capacity to conceptualize in the structure of human abilities? Is it the same thing as, or an aspect of, the general factor of intelligence? Is it linked in any way to the verbal factor and linguistic ability? Or to the spatial factor and the ability to appreciate patterns? Again, if rightly described by Price as a 'recognitional capacity', is it related somehow to some memory factor? There is little evidence at present enabling us to answer these questions, but they are questions of the kind that should prove amenable to psychological techniques.

Finally, the psychologist must map the course, and if possible elucidate the causes, of the breakdown of the capacity for conceptual thinking and its regression to lower levels in psychotic conditions, in old age, and as a result of brain injuries.

None of these tasks can be satisfactorily carried out without clear-cut operational definitions of conceptual capacities or without adequate tools. If the psychologist is interested in conceptual thinking as a capacity, the essential tool is one that measures this capacity—in other words a test or tests of known reliability and content. In view of the criticisms which have recently been levelled at tests as *ad hoc* instruments of no theoretical value it is perhaps important to insist that in research into human capacities they are the indispensable tools. There may be methodological problems in their use and interpretation (of which tools is this not true?), but they are not incapable of logical justification or scientific refinement. Existing tests of conceptual thinking, however, leave much to be desired. The Vigotsky test and the Goldstein-Scheerer sorting test have never been adequately standardized or analysed, and as Semeonoff & Laird (1952) point out in their recent study of the Vigotsky test it is practically impossible to demonstrate reliability. The main weakness both of Piaget's development studies, and the clinical studies of conceptual deterioration, is their lack of adequate and reliable measures. The first and most important task confronting the psychologist in the field of conceptual thinking is the construction of better tests. This is a prerequisite not only for research, but also for clinical application and diagnosis.

The practical value of investigations into conceptual thinking can hardly be doubted. The functioning of the normal adult mind is characteristically conceptual: our understanding both of the intellectual development of the child, and of the pathological regressions of later life depends upon our insight into the nature of conceptual thinking. Any techniques for assessing conceptual ability would have an immediate practical value for child and clinical psychologists. But over and above this the investigation of conceptual thinking should throw light on the nature of cerebral processes and of human intelligence, and assist in elucidating some of the central theoretical problems of psychology.

REFERENCES

AVELING, F. (1912). *On the Consciousness of the Universal and the Individual*. London.

COHEN, M. R. (1946). *A Preface to Logic*. London.

CRAIK, K. J. W. (1943). *The Nature of Explanation*. Cambridge.

FIELDS, P. E. (1932). The development of the concept of triangularity by the white rat. *Comp. Psychol. Monogr.* no. 42.

FISHER, S. C. (1916). The process of generalising abstraction and its product, the general concept. *Psychol. Monogr.* no. 90.

GOLDSTEIN, K. & SCHEERER, M. (1941). Abstract and concrete behaviour. *Psychol. Monogr.* no. 239.

HANFMANN, E. & KASANIN, J. (1942). *Conceptual Thinking in Schizophrenia.* New York.

HARLOW. H. F. (1949). The formation of learning sets. *Psychol. Rev.* LVI.

HAYEK, F. A. (1952). *The Sensory Order.* London.

HEBB, D. O. (1949). *The Organisation of Behaviour.* New York.

HEIDBREDER, E. F. (1946). The attainment of concepts. I–II. *J. Gen. Psychol.* XXXV.

HEIDBREDER, E. F. (1947–9). The attainment of concepts. III–VIII. *J. Psychol.* XXIV–XXVII.

HULL, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychol. Monogr.* no. 123.

HUMPHREY, G. (1951). *Thinking.* London.

JAMES, WM. (1890). *Principles of Psychology,* vol. I, p. 460.

LANGER, S. (1942). *Philosophy in a New Key.* Harvard.

LASHLEY, K. S. (1943). Nervous mechanisms in learning. In Murchison's *Handbook of Experimental Psychology.*

PAVLOV, I. P. (1927). *Conditioned Reflexes.* Engl. tr.

PIAGET, J. (1926). *The Language and Thought of the Child.* Engl. tr. London.

PIAGET, J. (1950). *The Psychology of Intelligence.* Engl. tr. London.

PRICE, H. H. (1953). *Thinking and Experience.* London.

REED, H. B. (1946). Factors influencing the learning and retention of concepts. I–V. *J. Exp. Psychol.* XXXVI–XL.

RIBOT, T. (1897). *L'evolution des idées générales.* Paris.

SEMEONOFF, B. & LAIRD, A. J. (1952). The Vigotsky test as a measure of intelligence. *Brit. J. Psychol.* XLIII.

VIGOTSKY, L. (1934). Thought in Schizophrenia. *Arch. Neurol Psychiat., Lond.,* XXXI.

WALTER, G. (1953). *The Living Brain,* p. 130. London.

WOLTERS, A. W. (1933). On conceptual thinking. *Brit. J. Psychol.* XXIV.

(*Manuscript received* 9 *October* 1953)

CHANGES OF PERFORMANCE WITH AGE ON SIMILAR TASKS OF VARYING COMPLEXITY

By HILARY M. CLAY

The Nuffield Research Unit into Problems of Ageing, The Psychological Laboratory, University of Cambridge

I. *Introduction* (p. 7). II. *Method* (pp. 7–8). III. *Results* (pp. 8–12): (i) *Accuracy* (pp. 8–10); (ii) *Number of alterations* (pp. 10–11; (iii) *Time taken* (pp. 11–12). IV. *Discussion of results* (pp. 12–13). V. *Summary* (p. 13). *References* (p. 13).

I. INTRODUCTION

Comparisons between young and old on tests of mental ability have shown a decline of performance with age. This deterioration was at first attributed solely to the time limits of the tests used. Lorge (1936) stated that 'speed obscures mental power' which 'probably does not deteriorate with age'. This view was supported by Miles (1938) with the reservation that 'older adults maintain their ability as long as they continue mental practice'. Lawton (1943) expressed his opinion more cautiously: 'If we eliminate speed and consider only mental power, the difference between older and younger individuals is reduced.'

Although the slowing of receptor and effector processes associated with age is the main cause of decline of performance on tests in which success depends on speed, an explanation is still to be sought for the differences observed between younger and older subjects on untimed tests (for example, Raven, 1948; Foulds & Raven, 1948). Other factors, such as length of problem and degree of familiarity with test material, have been examined, but unfortunately mainly in relation to speed. The present experiment is an attempt to isolate one factor, that of complexity in terms of the number of items with which a subject has to deal simultaneously but *at his own speed*, and to study the effect of varying complexity on ability to solve a problem task as one gets older.

II. METHOD

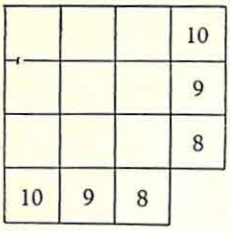
Description of task

Subjects were required to arrange counters, each stamped with a number from 1 to 4, in rows and columns of squares on a board so that the sum of the numbers on the counters in each row and each column equalled the marginal total printed for each on the board.

Four problems were used, varying in the number of counters given, with corresponding changes in the number and value of the rows and columns (see Fig. 1).

In each task the display was similar. The boards were made of white cardboard on which horizontal and vertical black lines were drawn 3·5 cm. apart. The white ivorine counters were 2·8 cm. in diameter, and on each was engraved in black a number 1·3 cm. high. They were stacked in a wooden rack according to their values.

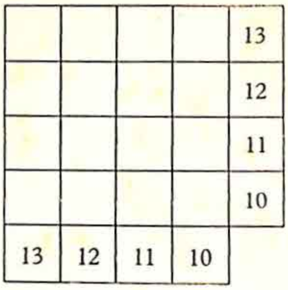
The instructions were identical for the four problems. The subjects were told that they should place one counter in each square and that the numbers on the counters should add up, both across and down, to the total given for each row and each column of squares.



(*a*) Problem 1

Counters:

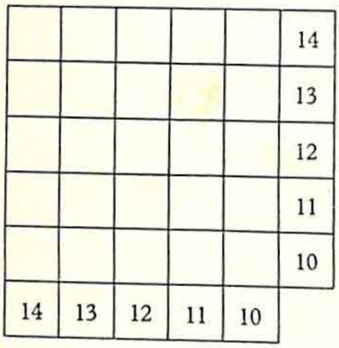
Value	1	2	3	4
Number	1	1	4	3



(*b*) Problem 2

Counters:

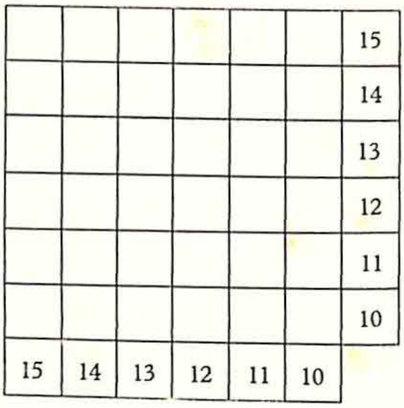
Value	1	2	3	4
Number	2	3	6	5



(*c*) Problem 3

Counters:

Value	1	2	3	4
Number	7	6	7	5



(*d*) Problem 4

Counters:

Value	1	2	3	4
Number	14	10	7	5

Fig. 1. The boards used in the four problems, showing numbers and values of marginal totals and of counters presented.

The subjects

One hundred and twenty-eight subjects were tested. They were divided into two groups according to age:

Younger group. Sixty-four naval ratings, mean age = 20·9 years (range 18–24 years).

Older group. Sixty-four pensioners, mean age = 67·5 years (range 55–78 years).

Both these groups were divided into four equal sub-groups, to each of which one of the four problems was presented.

All the older subjects, who volunteered to do the experiment, maintained a wide range of interests, although retired from active work.

III. Results

(i) *Accuracy*

The achievements of the older and younger groups were approximately equal on Problem 1. The number of subjects in both groups who solved the problem fell as the number of counters was increased (see Table 1). The fall was more marked for the older groups. The five older subjects who failed to solve Problem 2, nevertheless managed to get all but two of the marginal totals correct, and the differences between the younger and older groups on this problem and on Problem 3 were not significant. On Problem 4, however, the extent of the errors made by the older group was significantly greater than that of the younger ($P < 0{\cdot}05$).[1]

Table 1. *Distribution of subjects according to number of incorrect marginal totals*

		No. of incorrect totals										
Problem 1	0	2	3	4	5	6						
Younger group	15	1	—	—	—	—						
Older group	14	2	—	—	—	—						
Problem 2	0	2	3	4	5	6	7	8				
Younger group	15	1	—	—	—	—	—	—				
Older group	11	5	—	—	—	—	—	—				
Problem 3	0	2	3	4	5	6	7	8	9	10		
Younger group	11	2	1	2	—	—	—	—	—	—		
Older group	7	5	1	2	1	—	—	—	—	—		
Problem 4	0	2	3	4	5	6	7	8	9	10	11	12
Younger group	12	—	1	—	2	1	—	—	—	—	—	—
Older group	5	3	—	3	1	2	—	—	1	—	—	1

Table 2. *Analysis of errors*

	Total no. of subjects making errors	Type of error			
		Simple correction needed	Complex correction needed	Failure to place all counters	No attempt to correct
Problem 1					
Younger	1	1	—	—	—
Older	2	2	—	—	—
Problem 2					
Younger	1	1	—	—	—
Older	5	2	3	—	—
Problem 3					
Younger	5	1	4	—	—
Older	9	2	3	—	4
Problem 4					
Younger	4	—	3	1	—
Older	11	1	4	4	2
Totals					
Younger	11	3	7	1	—
Older	27	7	10	4	6

The errors made may be divided into four types:

(*a*) A simple error involving either two rows or two columns, which is easily corrected by a direct exchange of two counters. This error may be due to carelessness in checking the totals, but there was some evidence from the subjects' comments that it occurred because they did not realize that by altering the positions of two counters within, for instance, a row, they could correct an error in a column without altering the total of the row. As Table 2 shows, the older group made slightly more errors of this type than the younger.

(*b*) A more complex error involving two or more rows and columns, which requires not only the inter-change of counters in the incorrect rows and columns, but frequently a transference of the error to a correct row or column before a successful solution can be reached. No errors of this type appeared on Problem 1, and on Problem 2 they were made only by older subjects. Subjects in both groups made errors of this type in Problems 3 and 4.

(*c*) Failure to complete the task. This type of error only occurred in Problem 4. Four of the older subjects compared to one of the younger, failed to place all the counters on the board. They had managed to arrange some of the rows and columns correctly, and then realized that with the counters they had left it would be impossible to complete the problem without altering the positions of those already on the board. This they were unwilling to do.'

---

[1] The tests of significance used were those based on the ranking method developed by Kendall (1948).

(*d*) Although this type is not exactly an error, it is a failure to solve the problem by making no effort to correct mistakes. This type of response was made only by older subjects in Problems 3 and 4. These subjects knew they had made errors, but refused to attempt to correct them on the grounds that either it would take too long or it would upset the correct rows and columns.

(ii) *Number of Alterations*

The number of alterations made in the position of each counter *after* it had once been placed on the board was recorded. This measure gives some indication of attempts to correct errors and of flexibility of methods of tackling the task.

The only significant difference between the two groups in the number of alterations made was in Problem 3, where the younger subjects made more alterations than the older ($P < 0{\cdot}05$) (see Table 3).

Table 3. *Distribution of subjects according to number of alterations made, and the mean number of alterations for each group*

No. of alterations made	Problem 1		Problem 2		Problem 3		Problem 4	
	Young	Old	Young	Old	Young	Old	Young	Old
0	—	2	3	1	—	1	1	—
1–4	3	2	3	4	1	4	—	2
5–16	7	7	5	3	3	3	3	4
17–64	6	5	5	6	9	7	8	8
Over 64	—	—	—	2	3	1	4	2
Mean	12·4	12·6	13·8	25·9	38·2	20·1	37·1	38·7

As the number of alterations was closely linked with the way in which difficulties were overcome during the performance on the task, it is worth while considering the different methods employed:

(*a*) Perhaps the most drastic way of coping with difficulties was that adopted by six of the sixty-four younger subjects, and by two of the sixty-four older subjects. They removed all the counters from the board and started again. The two older subjects were in the group doing Problem 4, and one of them replaced the counters in almost identical positions the second time. He was not adopting a new method. On the other hand, the younger subjects took the opportunity of trying a new approach to the task.

(*b*) Another method involved the exchanging of all the counters in one row, for example, with all those in another, even if this meant that counters of the same value occupied identical positions. This was one way of overcoming the difficulty of correcting errors when no direct exchange of two counters was possible. Only three younger and four older subjects adopted this method.

(*c*) The most usual way of making alterations was by moving single counters at a time. Often subjects took a counter of the required value from another position on the board, regardless of the effect such a move would have on the solution as a whole. This method involved a number of useless alterations; as one subject remarked: 'All that juggling and it's still wrong.'

These methods suggest that some subjects were tackling the problem in a series of trial-and-error attempts. While it is true to say that a final solution was often reached after a number of piecemeal alterations had been made, an analysis of the initial methods adopted showed that, in the main, subjects approached the problem with some plan, however vague.

The older subjects chose methods similar to those of the younger, but were less successful in carrying them out as the number of counters increased. Fifteen of the sixty-four older subjects compared to three of the younger made an error in the arrangement of the first marginal total completed. Seven of these fifteen older subjects were doing Problem 4. On Problems 3 and 4, the older subjects who started with the intention of arranging the counters so that they added up in both rows and columns, soon forgot or neglected one dimension and concentrated on the other, without correcting their errors before they finished. Or, if they had adopted the method of arranging the counters along one dimension before attempting to fit in with the totals in the opposite direction, they would make the alterations necessary to correct for the other set of totals without any regard to the effect these made on the previously correct arrangement. They treated the problem as two separate tasks.

Table 4. *Distribution of subjects according to times taken, and the mean time for each group*

	Problem 1				Problem 2				Problem 3				Problem 4			
	Younger		Older		Younger		Older		Younger		Older		Younger		Older	
Seconds	C	I	C	I	C	I	C	I	C	I	C	I	C	I	C	I
Under 120	1	1	6	—	3	—	—	1	—	—	—	—	—	—	—	—
121– 240	7	—	2	—	5	1	2	1	2	—	—	2	—	—	—	—
241– 480	6	—	5	1	5	—	4	2	3	2	3	3	2	—	—	2
481– 960	1	—	1	1	2	—	—	1	2	—	4	1	3	3	5	2
961–1920	—	—	—	—	—	—	4	—	4	2	—	3	7	1	—	4
Over 1920	—	—	—	—	—	—	1	—	—	1	—	—	—	—	—	3
Mean time	253		241		256		681		887		579		1007		1166	

C = Correct solution. I = Incorrect solution.

The majority of the older subjects carried on a running commentary throughout, and resorted to such *aides-mémoires* as moving a counter slightly out of place or keeping one finger on the board to indicate an incorrect row or column, while they checked the rest. One or two of the younger subjects counted aloud, but only during the final check of marginal totals when they seemed to use verbal cues in much the same way as the older subjects made use of spatial ones.

(iii) *Time taken*

Although the subjects were told that there was no time limit, the total time spent on the task was recorded (see Table 4).

Some unpublished results by Kay indicate that speed of performance falls more strikingly with age at difficult than at easy tasks. On the other hand, Birren, Allen & Landau (1952) found that, under limited-time conditions, older subjects were slower in adding digits than younger subjects, and were relatively slowest for the very short or easy problems.

The present results, shown in Table 4, do not entirely agree with either of these previous findings. The average times spent on all the problems, except Problem 2, by younger and older groups were similar. The older subjects took significantly longer than the younger to complete Problem 2 ($P < 0{\cdot}05$).[1]

A possible reason for these results is found when time taken is considered in relation to difficulty and achievement. Most of the younger subjects on all the problems and of the older subjects on Problems 1 and 2 who took a relatively long time produced correct solutions. This is consistent with the well-known principle that as difficulty (due either to rising demands of the task or to falling capacity of the subject) increased, success can be maintained up to a point by the expenditure of more time. With the older subjects on Problems 3 and 4 this principle seems to break down, and most of those who take a long time fail to produce a correct solution. It would seem that age and the demands of the problems together have produced such a degree of difficulty that for many subjects no amount of extra time they are likely to be able to spend will achieve success. As a result they either plod on unsuccessfully for a long time, or quickly realize that the task is hopeless and give up.

IV. DISCUSSION OF RESULTS

The ability of older subjects to solve the problems is clearly related to the number of counters involved.

The older subjects were able to do the easiest problem, with nine counters, as well as the younger.

There was, however, a progressive breakdown of performance on the part of the older subjects, resulting from the increasing complexity:

(*a*) Sixteen counters: the older subjects required *more time* than the younger.

(*b*) Twenty-five counters: the older subjects were *less active* in correcting for errors.

(*c*) Thirty-six counters: the older subjects were *less accurate*, and 25% of the older group *failed to complete* the task.

Each of these differences was significant.

This breakdown cannot have been entirely due to the decline in accuracy with age in adding as shown by Birren & Botwinick (1951) and by Birren *et al.* (1952). The numbers of errors made in Problems 3 and 4 of this experiment are greater than the numbers made by their subjects adding five and six digits respectively. At the same time, it is clear that the errors remaining in the solutions were not mere errors of addition, because those which were made in the early stages of the task were later corrected—those which remained were recognized but were left because the subjects were unable to correct them.

It seems, therefore, that it was the increasing amount of material presented which caused older subjects more and more difficulty. They could readily *perceive* and *understand* what they were required to do. Failure appears to have been due to changes with age in the amount of data that could be organized together. They were unable to build a solution in a logical way with the more complex problems. Confusion was apparent. They could see their errors, but were unable to make the necessary connexions between errors to arrive at a correct solution. The younger subjects encountered similar difficulties in coping with the increased data, but were methodical in dealing with them. They followed their initial plan more closely, whereas the older subjects became muddled and then neglected some parts of the task.

[1] In a control experiment, no difference was found between the times taken by subjects of different age groups to place twenty-five blank counters in squares on a board.

As the complexity of the tasks was increased, the decline of ability with age was revealed in different ways. This serves to emphasize the necessity, which has been found in previous experiments on ageing (Welford *et al.* 1951), of considering several aspects of performance together when comparing different age groups.

It is perhaps tempting to see these results as linked with a failure of short-term memory, which would make it difficult to relate one piece of evidence in a complex pattern with another. This approach has already been suggested by Welford *et al.* (1951, p. 101) as an explanation of a different type of problem-solving experiment. They point out that the failure may not necessarily be of the memory processes themselves. Any general slowing would be likely to increase the load upon short-term memory, in the sense that whatever was carried would have to be carried for a longer period before it could be used.

V. SUMMARY

In this experiment an attempt was made to study the effects of changes in the number of items to be dealt with simultaneously but without time-stress on the ability of older compared to younger subjects to solve a problem task. Sixty-four subjects under 25 years of age and sixty-four subjects over 55 years of age were tested. They were required to arrange counters of different values in rows and columns on a board, in such a way that the sum of the counters equalled a marginal total given for each row and each column. Four problems were used, in which nine, sixteen, twenty-five and thirty-six counters respectively were presented. Each subject was given one problem only. The results showed that the performances of the two groups were similar on the simplest problem but, as the complexity of the task increased, the older subjects had greater difficulty compared with the younger.

I wish to thank the members of the Cadbury Pensioners' Club who acted as subjects, and also Cadbury Brothers Limited who kindly provided facilities for testing. The naval ratings took part in the experiment through the courtesy of the Admiralty.

My sincere thanks are due to Mr A. T. Welford, Director of the Unit, for his valuable criticism throughout.

REFERENCES

BIRREN, J. & BOTWINICK, J. (1951). Rate of addition as a function of difficulty and age. *Psychometrika*, II, 219–32.

BIRREN, J., ALLEN, W. R. & LANDAU, H. G. (1952). Speed, accuracy and difficulty. *Rep. psychometric lab. Univ. Chicago*, no. 76.

FOULDS, G. A. & RAVEN, J. C. (1948). Normal changes in the mental abilities of adults as age advances. *J. Ment. Sci.* XCIV, 197.

KENDALL, M. G. (1948). *Rank Correlation Methods*. London: Charles Griffin and Co.

LAWTON, G. (1943). Aging mental abilities and their preservation. In *New Goals for Old Age*, ed. by G. Lawton, pp. 11–33. New York: Columbia University Press.

LORGE, I. (1936). The influence of the test upon the nature of mental decline as a function of age. *J. Educ. Psychol.* XXVII, 100–10.

MILES, W. R. (1938). Psychological aspects of aging. In *Problems of Aging*, ed. by E. V. Cowdry. Baltimore: Williams and Wilkins.

RAVEN, J. C. (1948). The comparative assessment of mental ability. *Brit. J. Psychol.* XXIX, 12.

WELFORD, A. T. *et al.* (1951). *Skill and Age*. Oxford University Press for the Nuffield Foundation.

(*Manuscript received* 5 *May* 1953)

MEMORY MECHANISMS AND THE THEORY OF SCHEMATA[1]

By R. C. OLDFIELD

Department of Psychology, Reading

I. *Introduction* (pp. 14–15). II. *The character of recall* (pp. 15–16). III. *Circuital storage devices* (pp. 16–18). IV. *Memory systems based upon circuital storage elements* (pp. 18–19). V. *Five questions arising* (pp. 19–22). VI. *Summary* (p. 22). *References* (p. 23).

> The rudiments of memory are laid in the perpetual recurrency of the same impressions, and clusters of impressions. The traces which letters, and words, i.e. clusters of letters, leave afford an instance and example of this. And, as in languages, the letters are fewer than the syllables, the syllables than the words, and the words than the sentences, so the single sensory impressions and the small clusters of them, are comparatively few in respect of the large clusters; and, being so, they must recur more frequently, so as the sooner to beget those traces which I call the rudiments or elements of memory. When these traces or ideas begin to recur frequently, this also contributes to fix them, and their order, in the memory, in the same manner as the frequent impressions of the objects themselves.
>
> DAVID HARTLEY, *Observations on Man*, 1749

I. INTRODUCTION

The conception that, underlying the power of recalling past experiences, there exist *traces* in the mind or brain is not only a very old one, but by its simplicity gained wide acceptance among philosophers and psychologists until quite recently.[2] The idea that, as a necessary condition of reproduction, some *simulacrum* of the past occasion must be preserved carried with it the implication that the alterations in the matter of recall must be quantitative, comparable to the gradual erasure of an inscription on stone, or to the fading of a photograph. Thus many important features of recall were disregarded, though from a theoretical point of view a number of attempts were made to cope with the difficulties of 'fixed, lifeless traces'.[3] Many, however (especially physiologists and neurologists), who had occasion to make use of the conception showed themselves notably blind to the explanatory pitfalls entailed.[4]

The experimental work of Bartlett (1932), following that of Phillippe, disclosed a number of features in the process of recall which ill accorded with the doctrine of fixed, lifeless and independent traces. With some of these difficulties the Gestalt school, especially Koffka (1935), attempted to deal by a relaxation of the original doctrine of traces, in the sense of postulating dynamic interaction between them. Within certain relatively simple and delimited fields this extension of the doctrine may be said to have been heuristically valuable. But intolerable complexities and uncertainties arise when the attempt is made to extend the theory generally, as students of Koffka (1935) will note. Bartlett, seeking a fresh approach based upon the theoretical views of Head (1921),[5] interpreted his experimental findings in terms of a conception which he named the *schema*. This assuredly raises its own problems, but dispenses with the trace as the basic element in the mechanism of remembering.

---

[1] Read at meetings of the Northern Branch, British Psychological Society, 16 May 1953, and of the Bristol University Psychological Society, 8 May 1953.

[2] Gomulicki (1953) gives a useful account of the history of trace theories.

[3] For example, the 'apperceptive mass' of Wundt, and the 'dispositions' of Stout.

[4] Cf., for example, Munk (1890).

[5] For reviews of Bartlett's theory and its relation to that of Head, see Oldfield & Zangwill (1942–3), and Northway (1940). Brain (1950) has considered some wider applications of the theory.

The trace is a theoretical conception whose concrete exemplar may be any one of those more or less permanent modifications of material substance which, in some more or less direct sense, represent an object or occasion, and which play so large a part in our everyday existence. For all its limitations as a model, it is simple and, as a conception, easily handled in thought. If, however, it must be rejected on account of demonstrated impotence to order empirical findings, it may well be asked what shall replace it. Now it may be noted, that the corresponding method of storing information in terms of permanent traces is equally out of favour with designers of modern computing machines, although for different reasons. In this latter case, a quite novel principle of storage has been evolved. And suggestions have been often made that such a device might form a better model for human memory than does the older one based upon the photograph or gramophone record. Some colour is lent to proposals of this kind by the demonstration of structures in the central nervous system which incorporate inter-connexions similar in kind to those required in the memory-circuits of a modern computing machine. It should be noted, however, that the *known* functions of these structures relate to the temporal prolongation of inhibitory and excitatory effects. No direct connexion with the mechanism of memory as such has been demonstrated. It has been urged, too, that there are grave objections to such a view of the physiological basis of memory (e.g. Eccles, 1953).

The suggestion that the memory-circuit of a computing machine might afford a better analogy for the consideration of remembering has generally been made upon the basis that such a device possesses greater functional flexibility than does the semi-permanent trace. So far as I am aware, however, no effort has ever been made to work out the conception in more detail. In particular, it has not been asked whether a memory founded on this kind of basic element could go any further than does the trace in providing a theoretical formulation agreeable to empirical findings of the kind expounded by Bartlett and his followers. To make this effort is the object of this paper, but discussion will be confined to certain specific points, and no attempt will be made to formulate a general theory of remembering in these terms.

I will first very briefly recapitulate the relevant points of Bartlett's findings; and try to indicate the vital features of his theory. Secondly, the principle of the 'memory-circuit' as used in modern computing machines will be stated, and illustrated by a simple example. We shall then consider whether the latter can be regarded as a suitable element out of which a memory-mechanism could be built, and which would show at least some of the properties required by the known facts.

II. THE CHARACTER OF RECALL

Bartlett, convinced that the experimental constraints employed by Ebbinghaus in the interests of simplicity and quantification in fact prevented the emergence of the salient phenomena of recall, studied the recall of everyday material in the form of pictures and prose. The conditions in which the subject became acquainted with this material approximated to those of ordinary life. His conclusions were, briefly, as follows:

(1) Only in exceptional instances is recall *literal*. In general, changes of various kinds

are introduced into the material. One exception, in which recall may be literal, in effect exactly reproducing the original, is that of word-perfect recitation. Another, not mentioned by Bartlett, is that of immediate reproduction of a quantity of material which falls within the immediate memory span of the subject. In all other cases change is the rule.

(2) Such changes include not only omissions but also qualitative alterations, and even importations. Moreover, these changes are not arbitrary, but occur in accordance with the general principle that the material is converted into a form more in keeping with the cognitive conventions, logical and causal, which are current in the social *milieu* to which the individual belongs. There is *conventionalization* and *rationalization*. As Bartlett views it, the function of this kind of change is to render the material more easily handled in recall.

(3) There is a change in the relative emphasis of different parts of the material. Some elements, particularly details such as names, or other features which in the original are outstanding to the individual and attract his interest, possess especial survival-value. They may, indeed, be elaborated or exaggerated. These dominant details may be preserved in the form of a specific image, and they are thought by Bartlett to act, so to speak, as *labels* which identify the original experience, and as *starting-points* around which the rest of it may be rebuilt when the individual is faced with the need to recall.

(4) Recall is thus essentially *reconstructive*, not *reproductive*. Most of our experience is not retained as such, in explicit or concrete form, but is assimilated into the *schemata*, contributing its quota to the remoulding of their structure and adaptive capacity. Some fragments of experience, on the other hand, survive in a relatively unchanged and explicit form. Recall consists in reconstruction, based upon these latter, by the operation of the schemata, which themselves incorporate general laws and principles expressing the uniformities of experience.

Such a view, it is clear, suggests many special problems. But it is at once evident that, if we seek to interpret it in accordance with a mechanical analogy (and the trace-theory itself is a form of mechanical analogy), one important question will arise. How can information be stored in some general or abstracted form, and still allow some explicit, if incomplete, reconstruction of the original? It is sometimes suggested that *no* form of mechanical device could achieve this. But before this can be accepted more detailed consideration of the question is desirable. Accordingly, we shall state and illustrate the general principle according to which the memory-circuit of a modern computing machine operates. We shall then consider the extent to which a system built of such elements could be expected to store information in a generalized form, and subsequently reconstruct the original messages fed into it. It should perhaps be emphasized that such an attempt, though made in specific terms, can do no more than elucidate the basic requirements of the problem. It can, in itself, offer no direct suggestions as to the actual mechanisms, physiological or otherwise, which are involved in memory.

III. CIRCUITAL STORAGE DEVICES

The general principle of the storage devices used in modern computing machines is most simply stated by describing an actual instance. The first requirement is that the information to be stored should be suitably coded. We may, without loss of generality,[1] suppose that all messages consist of sequences of the scale-of-two digits 0 and 1. In this form the message can be propagated in an electrical circuit as a sequence of impulses. So, for instance, the message 101100011011 will appear as the wave shown in Fig. 1.

---

[1] It can be shown that anything that is 'information', in the sense that it can be explicitly formulated and represented in symbolic terms, can always be re-coded completely in this form.

Now consider Fig. 2. *AB* is a tube filled with mercury. At the end *A* is a transducer element which converts electrical into mechanical changes. These are propagated down the mercury and reach *B*. Here is situated a receptor element which converts them back into electrical changes. The electrical message-sequence from *B* can now pass along either or both of two routes. In the first it passes to, and may be used in, other parts of the apparatus. In the second it passes back to *A*, to be retransmitted down the tube. Deterioration in the form of the impulses, occasioned by the transmission and conversions, is corrected in the unit *X*, which also makes good incidental energy losses in the system.

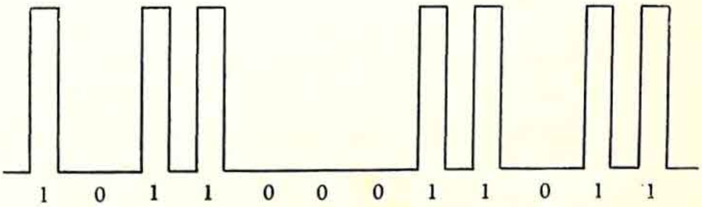


Fig. 1. Pulse-coded version of message 101100011011.

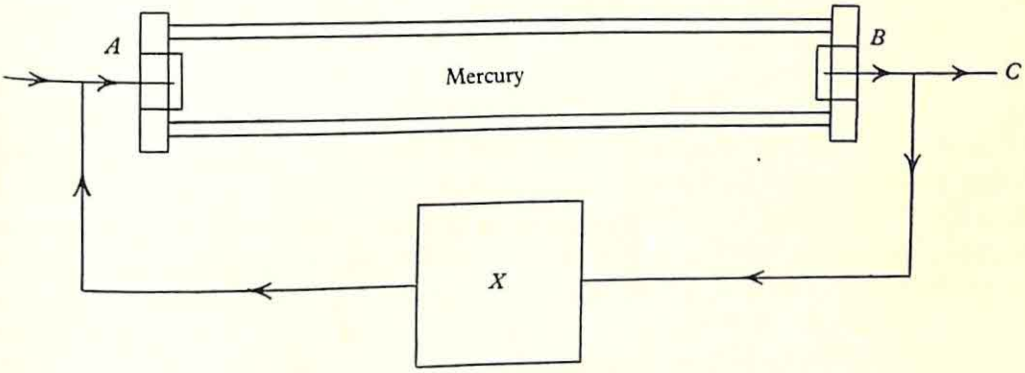


Fig. 2. Mercury circuital storage element.

If, now, we pass into the system a given message, it will circulate indefinitely, and is effectively stored. If the electrical circuit *BXA* is broken for a period corresponding to the time taken for the message to pass down the tube, the system is completely 'cleared', and can be used again to store another message. If the information contained in the message is to be made use of elsewhere in the apparatus, it can be passed out along *BC*, and can be used any number of times in this way. The length of the message which can thus be stored is approximately equal to the number of digits in a sequence which occupies the time taken for one impulse to pass down the tube. The time of transmission in the electrical parts of the circuit may be regarded as negligible.

The general principle of storage is thus a novel one and possesses several outstanding features. Among these are (1) the ease with which the message can be constantly repaired, so that it can be preserved unchanged over long periods of time; (2) the ease with which it can be annulled and wholly disposed of when it is no longer required; and (3) the ease with which it can be made available to other parts of the apparatus when required without affecting its continued storage. The particular device just described is but one of many possible ones operating upon the same circuital principle. It happens to be a relatively poor version so far as storage capacity is concerned, and although others have been developed which are much more favourable in this respect, economy in such devices is of basic importance, and is not irrelevant to our further discussion.

IV. MEMORY SYSTEMS BASED UPON CIRCUITAL STORAGE ELEMENTS

So far our discussion offers nothing novel as an analogue for the mechanism of recall. For if we merely postulate a number of independent circuital storage-elements, we have only returned to another, if perhaps more entertaining and recondite, form of the hypothesis of fixed lifeless traces. The question to be discussed is whether such storage elements could be imagined as built into a system in which there would be some economy of storage without loss of the power to reconstitute the original messages. It does not seem impossible to see, broadly, how this might be done. Consideration of this question raises a number of further points, admittedly speculative, which are of some interest in wider connexions.

We have a 'black box' capable of taking in messages in the form of sequences of 0's and 1's, and storing them. At some future time, on receipt of an appropriate stimulus, the box is to redeliver the original message. For a start we shall suppose that it is to redeliver it in exactly the original form—though we shall later consider relaxation of this condition. The question of what is meant by an 'appropriate' stimulus, and of how it operates to pick out the correct message stored in the box, is, of course, one of great interest and importance. But we shall not, in the first instance, try to answer it. All we wish to do at this stage is to demonstrate the possibilities of economical storage and of reconstruction.

It seems fairly clear that unless the various messages received by the box incorporate *some* detectable elements of common pattern, there is no possibility of economy in their storage. If, however, *any* common patterns run through them, there arises the possibility of re-coding them in briefer form, and so of storing more messages in a given storage space. *One* simple way in which community of pattern might manifest itself, which we may take as an example, is the following. Suppose that all messages handled by the box consist of thirty-six digits, and that the following is one of these:

100100011011100011100100011100100011.

On inspection, it can be seen that if broken up into sub-sequences of three digits each, all of them are of two kinds only, namely 100 and 011:

100.100.011.011.100.011.100.100.011.100.100.011.

If we make 1 stand for 100, and 0 stand for 011, we can re-code the message completely as follows:

110010110110.

In fact, in this particular case, further re-coding is possible in sequences of three, the only patterns in this case being 110 and 010, and the second re-coded version, if 110 = 1 and 010 = 0 is

1011.

Neglecting, however, for the moment, this possibility of a second re-coding, we may consider how the first re-coding can introduce economy of storage. Suppose that we store the sub-sequences 100 and 011 in two storage elements, and the re-coded version in a third. The original message can now be reproduced by causing this latter element to draw out the sub-sequences in appropriate order as given by the sequence of digits in the re-coded version. Thus, when the output from the third tube is 1, the sequence 100 is fed out of the box, and when the output of the third tube is 0, 011. The original message can, in principle, therefore, be reconstituted. Now the re-coded message is only a third as long as the original, and requires only a third of the storage space. If, therefore, one storage element was just able to store one *original* message, we can store three re-coded versions in the same space. For nine messages, all susceptible of re-coding on the same basis, only five storage elements are required—three for the re-coded versions, and two for the sub-sequences, in place of nine. It is clear that the advantage will be the greater, the greater the amount of common pattern manifested, and the larger the number of messages possessing this community.

Before we proceed any further a few subsidiary points should be cleared up. The digits 0 and 1 only allow for the re-coding of two types of sub-sequence. If there are more than two, scale-of-two numbers containing more than one digit will have to be used. Thus, if the sub-sequences 1001, 0101, and 0010 are found to constitute all messages, code numbers of two digits will have to be used for each, for instance 10, 11 and 01. Hence twice as much storage space will be required for the re-coded versions as in the case where only two types are present. Secondly, if a number of re-coded versions are to be stored in a single tube, some indication must be provided to serve as a full stop at the end of each message. This can be any sequence not used as a code for any of the sub-sequences. This circumstance, again, makes some slight extra demand on storage space.

On assumptions such as these it is possible to derive an expression for the economy of storage. This would, however, only be valid for the rather special type of pattern assumed in the incoming messages, and its actual form would be of less interest than is the general principle that economy can be effected without impairing accuracy of reproduction. A number of further questions now require consideration.

V. FIVE QUESTIONS ARISING

(1) As we have already remarked, re-coding with consequent economy is only possible if there are some recurrences, uniformities or common patterns in the incoming messages. This is the case in the information from the environment received by the human organism. On the other hand, it is not, of course, suggested that the type of pattern is that very simple one suggested above. By comparison with this latter, the patterns of human experience are complex and subtle. They may only reveal themselves when some amount of abstraction and elimination of irrelevant detail has been carried out, which would correspond with a re-coding process in our analogy. Thus, for instance, the messages 100111011101100110011101 and 110001001110110001 do not present any obvious similarity of pattern until both have been suitably re-coded, by putting 1001 = 1 and 1101 = 0 in the first message, and 110 = 1 and 001 = 0 in the second, when the same re-coded version, 100110, is obtained.

(2) How would the performance of such a mechanism as we have considered compare with the use of Bartlett's schemata in recall? Suppose that the box has had considerable 'experience' of incoming messages, and has acquired a store of type sub-sequences.

These represent common elements in the events in its 'environment'. Treated as a total organized system, which they form in virtue of the various connexions between them which could be generated by re-codings of higher orders, they might be said to form its 'schemata'. For it is upon the basis of them that reconstruction of particular past messages is possible. Furthermore, we may suppose that, once acquired, the system is used to filter fresh incoming messages in accordance with the hierarchical organization which it possesses. Thereby each new message could be re-coded in such a way that only a certain minimum of storage space is occupied. This minimum is that needed to represent the individual features of the message which cannot be subsumed in more general terms derived from the past experience of the box. Some messages will be banal, in the sense that they incorporate little that cannot be subsumed under the schemata. These will demand little storage space, and will induce little reorganization of the schemata. Viewed in terms of the schemata they may be said to be highly probable, and to contain but little information. Other messages, on the other hand, will present novelty, be improbable, and contain much information. These will only to a small degree be capable of reduction and subsumption by the prevailing schemata. They will demand relatively large storage space. On the other hand, they may induce the operation of fresh coding principles, and thus lead to reorganization of the schemata. In this sense it might be said that the box's schemata form an active organization of past experiences. The demands of novel messages of high information-content might be supposed to maintain within the schemata persistently active searching for fresh ways of reorganizing the existing system by exploration of the various alternative possibilities of re-coding. In this way complex, clumsy ways of storing information might, sometimes suddenly, be replaced by much simpler and more economical modes of coding and handling the same information. Perhaps in somewhat similar fashion 'insight' may bring about widespread reorganization of a human individual's response-tendencies.

(3) We have spoken, in this last paragraph, of two supposed processes which require some further comment. These are first the 'filtering' of fresh incoming messages through storage elements which contain type-subsequences, and secondly, a spontaneous activity of searching for fresh methods of re-coding material already stored, whether type-sequences or abbreviated, i.e. re-coded, versions of messages.

The filtering process is, in principle, not difficult to conceive, since type-sequences already stored are available for comparison with fresh material, without jeopardizing their continued storage. To take a very simple instance, suppose that the box contains only two stored type-sequences, 010 and 110, and that a fresh message,

010010110010110110,

is received and stored temporarily as it stands. The output from this storage element is fed to two discriminator units, to one of which is also fed 010, and to the other 110. When the two messages fed to a discriminator unit are identical it responds by sending out a characteristic signal, say 1 for the 010 unit, and 0 for the 110 unit. If the messages are not identical, then the unit does not respond. A sequence of digits is therefore formed by the combined output signals of the two discriminators, in this case 110100, is stored in a further storage element and constitutes the re-coded version. Once the re-coding is complete, the units used for the temporary storage of the original message are cleared and made available for use again.

As to the spontaneous activity of searching for fresh methods of re-coding material already stored, this might be conceived in the following terms. Given a sufficient number of spare storage units for holding temporarily the results of its operation, there is no particular difficulty in supposing the mechanism to explore all the possibilities of re-coding, since this is a type of operation familiar in computing machines. But the mechanism must be controlled in such activities by some general principle, and it must have some criterion according to which it accepts or rejects the results of operations, and continues or ceases its searching activities. Such a general principle might be that of tending to reduce the total amount of storage space occupied by the material stored to date. If this were reduced by any set of re-coding operations, the results would be accepted and stored; and older, now redundant, material annulled. Further activity along these lines would continue until no further economy of storage was produced.

(4) We have suggested that re-coding activities are dependent upon the possession of stored type-sequences, and it might well be asked how, since these must themselves be derived from messages received by the machine, the latter 'gets a start', so to speak, in laying down its 'schemata'. It may, perhaps, be as well not to broach the discussion of this problem, which bears a considerable resemblance to a more famous, but equally troublesome, psychological puzzle. We may remark, however, that if the memory-mechanism we are considering forms part of a more complex machine which interacts with its environment, it may be reasonably supposed that it is endowed with certain 'innate' simple type-sequences.

(5) So far we have been concerned principally to indicate that memory-mechanisms incorporating circuital storage elements could, in theory, effect economy in the storage of incoming information, and that such economy would take the form of storing general patterns common to a number of messages separately from the special features peculiar to each message. A somewhat similar principle would seem, to judge by the experimental findings of Bartlett and his followers, to operate in the mechanisms of human memory, but in this case a further notable feature is the imperfection of the process, as judged against the standard of literal recall. Reproduction, as we have seen, is rarely literal, and we should consider to what this feature would correspond in the model memory-mechanism we have outlined.

The 'defects' in the products of human memory consist in omissions, alterations, and importations. These changes take place in accordance with the principles of conventionalization and rationalization. The stage in the whole process at which these changes occur has been the subject of some dispute. Bartlett, it would seem (and also the Gestalt school), thought of them as being brought about chiefly as a result of reorganization of material during the period between the original presentation and the act of reproduction. That the act of reproduction itself introduces changes cannot be denied, but these have often been regarded as taking an accidental form, rather than as representing the activities of the schemata. The possibility that an important and characteristic part of the change occurs at the time of presentation, and during the immediate process of storage, has been closely considered by Gomulicki (1952) in a number of pertinent investigations. His conclusion is that a very considerable amount of 'editing' takes place at this stage, so that what is stored has the character of a précis of the original material.

Now consider the situation of the mechanism in the face of a message which does not entirely conform to the patterns already incorporated in the machine's 'schemata'.

RESEARCH LIBRARY

We might suppose it to act in one of three ways. In the first place, it might simply not store the offending material at all. Secondly, it might condone small departures from pre-existing patterns. Thirdly, it might store those parts of the message which do not conform 'in clear', while re-coding those parts which do. The first alternative is not of interest to us. The second would involve the possession of a criterion by the sequence-discriminators, according to which they would treat as an instance of a type-sequence any sequence which did not depart from type by more than a certain number of digits. So, for instance, if the criterion was 'not more than one digit wrong', and the existing type-sequences were 11001 and 00010, the message

110010101000010110111100100010

will be re-coded 100110, the offending sequences, namely 01010 and 11011, differing from the types by only one digit each, and being coded and stored as the types themselves. If reproduction of this message is called for, it will, of course, appear in a form different from the original, and the alterations could reasonably be said to be ones of 'conventionalization' inasmuch as the material is brought into a form which conforms to the machine's 'past experience' in general.

The third alternative way of handling non-conformatory parts of the material was, it was suggested, by storing such parts in their original form, in much the same way as users of some codes must sometimes leave 'in clear' unusual words not provided for in the code-book. It would be necessary in this case to enclose the sequence in question in a pair of (coded) parentheses to indicate that this material is in clear, so that in the reproductive process such sequences are run straight off without being decoded. If sequences which cannot be assimilated are treated in this way, some points of interest arise. In the first place, the material in question is preserved unchanged and is not subject to processes of 'conventionalization' which may occur in the processes of coding and decoding. Secondly, such outstanding features could act as 'labels', by which such messages might be identified in the course of the further transactions of the system. It seems possible, for instance, that such a label might be operative in the appropriate selection of the message when there is a demand for reproduction. In these respects it might be said that sequences which cannot be assimilated into the schemata resemble the dominant details and images which, in Bartlett's formulation are carried along relatively unchanged, and seem to afford the starting-point for reconstruction.

VI. SUMMARY

An attempt is made to decide whether the memory-model suggested by circuital storage devices in modern computing machines[1] is better adapted to meet the empirical data on recall brought to light by Bartlett and others, than is the type of model which postulates storage in the form of *simulacra*, or traces impressed upon a medium. It is pointed out

---

[1] It has been pointed out to me, and it should be emphasized, that the considerations of this paper are by no means dependent upon the *sequential* aspect of processes of storing information encountered in modern computing machines. A type of machine which stored incoming messages on punched cards, automatically searched through these and punched re-coded versions on fresh cards, rejecting the original cards into a waste-paper basket, would do as well in principle as an exemplar. In fact, this paper might, instead of having its present title, have been called 'Sketch-design for an automatic filing system incorporating a principle of storage-space minimization'. The mechanical details of such a model would, of course, be more difficult to imagine, and for this reason the electrical sequential processes we have considered afford a more helpful picture.

that unless the circuital storage-elements are functionally organized in such a way as to allow of re-coding of the stored messages, they offer in themselves no advance upon the ordinary trace hypothesis. However, the flexibility of the circuital device might be used to achieve systematic re-coding upon the basis of common patterns in the incoming messages. These common patterns would then be stored separately from the particular arrangements of them which constitute different messages. Not only could there be economy of storage space, but the original messages could be reproduced in their original forms. Furthermore, when we consider how such a mechanism might handle messages which contain patterns not already held stored, possible analogies to the processes of rationalization and conventionalization, and to the part played by 'dominant detail', emerge. It is suggested that such a system would have some of the properties of schemata as postulated by Bartlett.

Although such a model is not immediately applicable to laboratory experimentation upon memory processes, it might prove useful in re-orientating our conception of memory-mechanisms over a wide field, including problems of recognition and 'set', of perceptual equivalence, of retroactive inhibition and immediate memory span, and of amnesias and paramnesias.

REFERENCES

Bartlett, F. C. (1932). *Remembering*. Cambridge University Press.

Brain, W. R. (1950). The concept of the schema in neurology and psychiatry. In *Perspectives in Neuro-psychiatry*, ed. Richter. London.

Eccles, J. C. (1953). *The Neurophysiological Basis of Mind*. Oxford University Press.

Gomulicki, B. (1952). Recall as an abstractive process. Thesis submitted for the degree of D.Phil., Oxford.

Gomulicki, B. (1953). The development and present status of the trace theory of memory. *Brit. J. Psychol. (Monogr. Suppl.)*, no. 29.

Head, H. (1921). *Studies in Neurology*, II. Oxford University Press.

Koffka, K. (1935). *Principles of Gestalt Psychology*. London.

Munk, H. (1890). *Über die Functionen der Grosshirnrinde*. 2 Aufl. Berlin.

Northway, M. L. (1940). The concept of the 'Schema'. *Brit. J. Psychol.* xxx, 316–25, xxxi, 22–36.

Oldfield, R. C. & Zangwill, O. L. (1942–3). Head's concept of the schema and its application in contemporary British psychology. *Brit. J. Psychol.* xxxii, 267–86, xxxiii, 58–64, 113–29, 143–9.

(*Manuscript received* 18 *March* 1953)

THE CONTINUOUS REGULATION OF SKILLED RESPONSE BY KINAESTHETIC FEED BACK

BY C. B. GIBBS

Medical Research Council, Applied Psychology Research Unit, Cambridge

I. *Introduction* (pp. 24–25). II. *Apparatus* (pp. 25–27). III. *Procedure* (pp. 27–28). IV. *Results* (pp. 28–31). V. *The Tabetic Patient* (pp. 32–34). VI. *Discussion* (pp. 34–38). VII. *Summary* (pp. 38–39). *References* (p. 39).

I. INTRODUCTION

The dominant characteristics of skilled movement provide two main hypotheses of the principle or organization which has evolved to control movement and to make effective use of sense data from peripheral nerve endings.

A skilled movement of short duration may be continuously regulated by sense data which are always being fed back from a limb, so that the organism is actuated by the difference between the present state of the limb in motion and the intended final position of the limb. The principle of control would then be continuous 'error actuation' and there would be a direct analogy with a feed-back servo mechanism.

Alternatively, a skilled movement may be 'triggered off' as a unit in response to an external stimulus, but be independent of current sensory control during the initial phase of movement (Woodworth, 1899). It is evident that this phase cannot be continuously monitored by visual feed back in many cases of rapid movement (cf. Fitts, 1951), and Chernikoff & Taylor (1952) have recently confirmed that kinaesthetic reaction time is little less than that in other modalities. Following an argument of Craik's (1947), they claim that this latency would not permit the arrest, or continuous regulation of movements of short duration by the proprioceptive data which are fed to the higher motor centres. A similar view is expressed by Lashley (1951). On this view the selection of an appropriate response is uniquely determined by 'previous experience' and the nature of the external stimulus.

The distinction between the two forms of organization is fundamental, and may best be illustrated by comparing the control principle of a guided, with a projected missile. There is continuous feed back in the former case, and successive modulations of flight behaviour are possible; but in the latter case only one control decision precedes the triggered event, and there is no feed back to permit modification of its flight. Applied to biological mechanisms, the principle of continuous actuation by error would mean that learned responses could be rapidly modified to suit specific environmental conditions, whereas the theory of impulse discharge implies that the main effect of learning is so to 'calibrate' or condition the organism, that a familiar external stimulus can release a relatively automatic and stereotyped response. The survival of organisms has always depended on a good balance between versatility, and the quick action of a ready made response. It is possible therefore that during evolution a compromise has been effected between the principles although one or the other may be dominant in specific conditions.

From the physiological viewpoint, the concept of continuous regulatory feed back from the proprioceptors, has now attained respectable antiquity, since Bell (1826) first discovered the circle of nerves. Experiments on the transfer of high-grade skill (Gibbs, 1951) show that skilled responses are rapidly adapted to new environmental conditions which necessitate a change in the force or extent of a previously learned response. The principle of continuous error actuation, and the relevance of the terminology and physical models of servo engineering are not, so far, generally accepted. The difficulty is pointed out by Ruch (1951): 'the protagonists of feed back have been unable to envisage the precise manner in which proprioceptive information is used in the control of movement'.

The main purpose of the present paper is to present a theory of this kind which must be consistent with the main experimental results concerning skilled movement, and account reasonably for the serious objections which have been made to the notion of feed back. Some contemporary data and views are therefore briefly reviewed, and original experiments are described, which were designed to test some critical points at issue; in particular, it is shown that controlled changes in proprioceptive stimulation lead to significant changes in psychomotor performance, and the amount of positive transfer between two skilled tasks is determined largely by the relative adequacy of the proprioceptive data in the first and second tasks.

It is thought that these results can be reconciled only with the theory that skilled responses are continuously regulated by proprioceptive feed back.

II. APPARATUS

(a) *General task.* The apparatus set the operators a task of manual tracking with velocity control. A course-setting mechanism, with two rotating cams, caused a spot of light to wander in two dimensions across the screen of a cathode-ray oscilloscope. The 'Standard' course varied the velocity of the light continuously in two dimensions. The traverse cam, that moved the light from left to right, could be changed to give a more 'difficult' course. The courses are plotted in Fig. 1 and show the movement of the eyes in degrees per second when fixated on the light. Eighty seconds were required for each revolution of the cams; the course was then repetitive, and a standard period of practice was 11 revolutions, i.e. about 15 min. for all operators.

The subjects sat directly facing the screen of the C.R.O., and were 14 in. from it. They held in both hands at approximately elbow height a joystick that could be deflected in any direction about a universal coupling. Elbow supports were provided. The lever was mounted horizontally, pointing toward the operator.

(b) *In the free-moving conditions*, a hand movement to the right or left, up or down, moved the spot of light in the same plane and direction as the hand. The task was that of compensatory tracking, and the subjects tried to hold the light stationary on a fixed cross-wire and central circle of the C.R.O. against the off-setting action of the course-setting mechanism. Any spot movements were errors, and these deviations were summed separately by four counters, which were driven by the central member of the differentials (Fig. 2) and showed errors of 'high', 'low', 'right' or 'left'. The errors were noted at each revolution of the cams and the counters recorded one unit per 0·05 in. of spot movement. The diagrammatical layout of the gear is shown in Fig. 2.

Displacement/velocity or rate control was used and each separate angular position of the joystick resulted in a different velocity and direction of movement of the spot of light. The length of the joystick and the amplification of signal could be altered so that different 'gear ratios' could be obtained between the amount of hand movement and the velocity given to the light. The gear ratio (i.e. stick sensitivity) remained approximately constant over the whole range of joystick movement when the sensitivity was set at any given value. Several gear ratios were compared to determine the optimal range of this physical variable and a standard ratio of spot velocity per 1 mm. movement of the hand was selected as representing the optimal movement conditions. This standard ratio was used

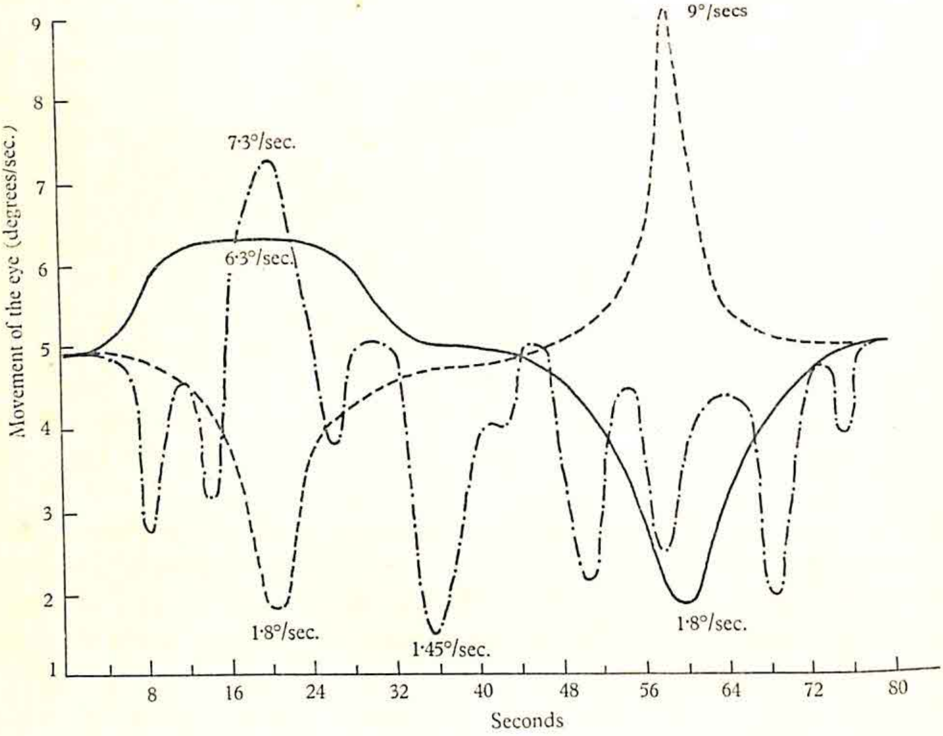


Fig. 1. The courses. Degrees per second at the eye plotted against time. – – – –, standard elevation course; ——, standard traverse course; –·–·–·, difficult traverse course.

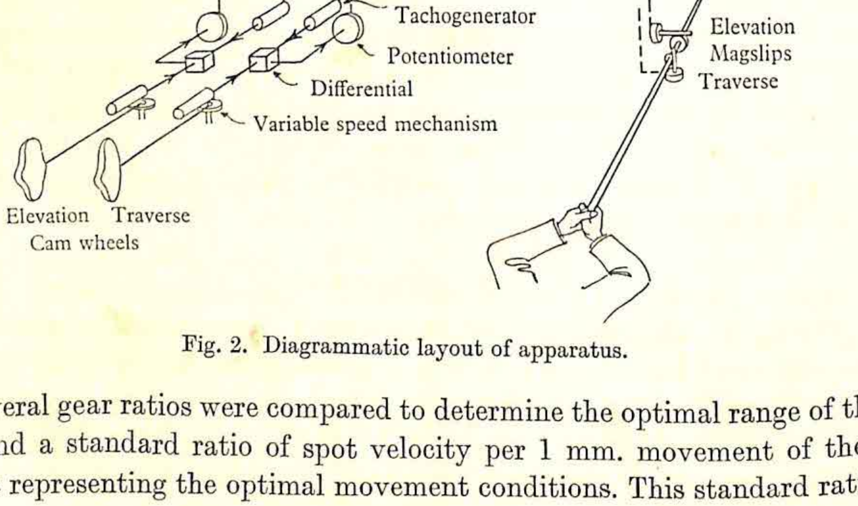


Fig. 2. Diagrammatic layout of apparatus.

in comparing the methods of free-moving and pressure control; 1 mm. of hand movement altered the velocity of the spot of light, subtended at the eye, by 0·53°/sec. The length of the lever was 19·5 in. from fulcrum to hand grip, and it was counterbalanced.

(c) *In the pressure control arrangement* of the apparatus, the situation was exactly as before except that the joystick was secured at the end remote from the operator and in the position of zero output. Applied pressure would then cause slight deflexions of the lever with spring loading tending to return the lever to centre. The joystick deflexions were converted and amplified, and used to control the movements of the spot, as in the free-moving arrangement. The amplifiers were used in both the pressure and free-moving conditions, but the gain was 17 times greater in pressure control. Variations in applied pressure therefore altered the velocity of the spot, and the pressure/velocity gear ratio was approximately constant over the range of pressure that was used, for each setting of the amplifier.

A number of different pressure/velocity ratios were compared by altering the spring loading and a standard ratio was selected and used in comparing the methods of free-moving and pressure control. In these standard conditions an application of 1 oz. pressure at the lever altered the velocity of the light by 0·26°/sec. subtended at the eye. The length of the lever was 19·5 in.

(d) *The time constant of the control system* is an important variable with considerable effects upon human performance. The time constant during the present experiments was 0·08 sec., i.e. there was a time lag between moving the control lever, and the instant when the spot of light reached its new steady speed, and this would be 0·08 sec., if the initial rate of change were maintained.

(e) *Finger operated controls.* The control lever was rotated upwards through an angle of 45° and then operated by the fingers and wrists of both hands, with supported wrists. The main comparisons of the finger controls were made with unchanged gear ratios at the controls (0·53°/sec. per 1 mm., and 0·26°/sec. per 1 oz.), since these ratios proved suitable for finger as for forearm controls.

(f) *The single-correction series of experiments* was designed to compare the times taken to correct a given amount of error at the screen, using the standard pressure and standard free-moving methods of control. Small target spots were painted on the screen, 40 mm. distant from the centre. The task was to move the spot of light from the paint spots, to within a small circle of 0·1 in. diameter at the centre of the screen, with the highest possible speed and to a specified accuracy. The course-setting mechanism was running during the tests and the spot tended to wander at a fixed rate, so that the operators had the dual task of (i) making a large corrective movement, and (ii) compensating for a constant course input. Graphic records were taken of the periods between (i) the starting signal to the subject, (ii) the initial movement of the spot of light, and (iii) the instant when the spot of light was arrested within the circle.

IV. PROCEDURE

Subjects. The ninety-five subjects used in the main experiments were Naval ratings aged between 19 and 26 years. About 80% of the subjects had previous experience of visuo-motor skills such as machine operation which were related in some way to the experimental task.

Practice effects. Practice effects were equalized for the test groups by presenting the conditions in Latin square form. But when only two conditions were compared, two groups of matched subjects were tested after equal training. The subjects were given 80 sec. of preliminary practice on the free-moving control, and the scores were used to divide the subjects into two groups, A and B, which covered the same range of initial ability.

All subjects had a minimum of 2 hr. preliminary practice before any comparisons were made, excepting the thirty-six subjects (groups A and B) who tested initial learning and transfer effects at the two controls.

V. RESULTS

(*a*) *The comparison of the sensitivities*

Meaningful comparisons could not be made between the controls without preliminary tests to find the optimal conditions for each. Nine subjects therefore tested a range of different sensitivities at the pressure control as shown by Fig. 4 A after equal practice on each sensitivity. The same subjects had 3 further days' training on the free-moving control before testing a corresponding range of sensitivities at this control (Fig. 4 B). The plotted points show the average total errors for the groups, taken over a full practice period of 15 min.

Accuracy improved nearly linearly at the pressure control as the pressure range was progressively increased, until an optimal point was reached. There was a similar effect at the free-moving control as the rate of movement was increased. After reaching the optimal sensitivity, there was a decrease in tracking accuracy when the pressure and movement range respectively were further increased at the two controls.

Pressure and free-moving gear ratios were selected as the standard conditions on the basis of these results, and they were always used in the subsequent tests of the controls.

There was some small movement at the *pressure control lever* and a total movement of about 0·08 in. was required to vary output velocities between 0° and 9°/sec.

(*b*) *Initial learning and transfer effects*

Thirty-six other subjects were divided into two groups. Groups A and B were matched in initial ability as previously described. An analysis of covariance was performed on the scores obtained in preliminary practice and on the 3rd and 6th day of practice, and it showed that the initial scores had predicted the later scores; thus final performance was dependent on initial ability. The analysis also showed that there were remaining differences between pressure and free-moving control, which could not be attributed to differences in the initial ability of the subjects.

The initial learning and transfer effects for both groups of subjects are shown in Fig. 3 with error plotted against time of practice. The points first plotted on the graphs show the total average errors for each group during the first 80 sec. of practice. The other plotted points show the average group errors per 80 sec. over a complete practice period of 15 min.

Group A first practised the pressure control for 3 days and were consistently better during this period than group B, who first practised the free-moving control. The groups exchanged tasks on the 4th day and the errors of group A were appreciably increased after transfer from the pressure control to the more difficult free-moving control. But in

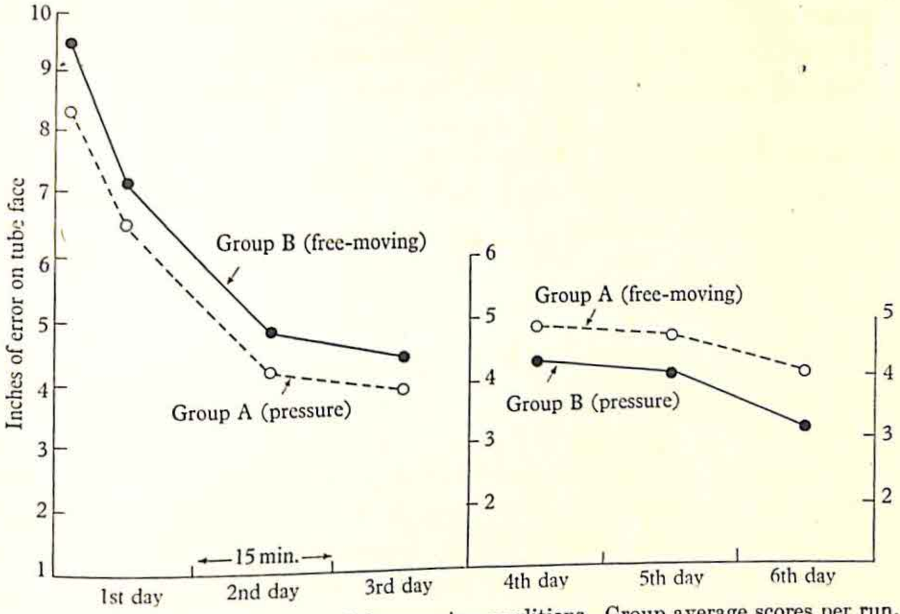


Fig. 3. Learning curves in pressure and free-moving conditions. Group average scores per run. - - - -, group A; ——, group B.

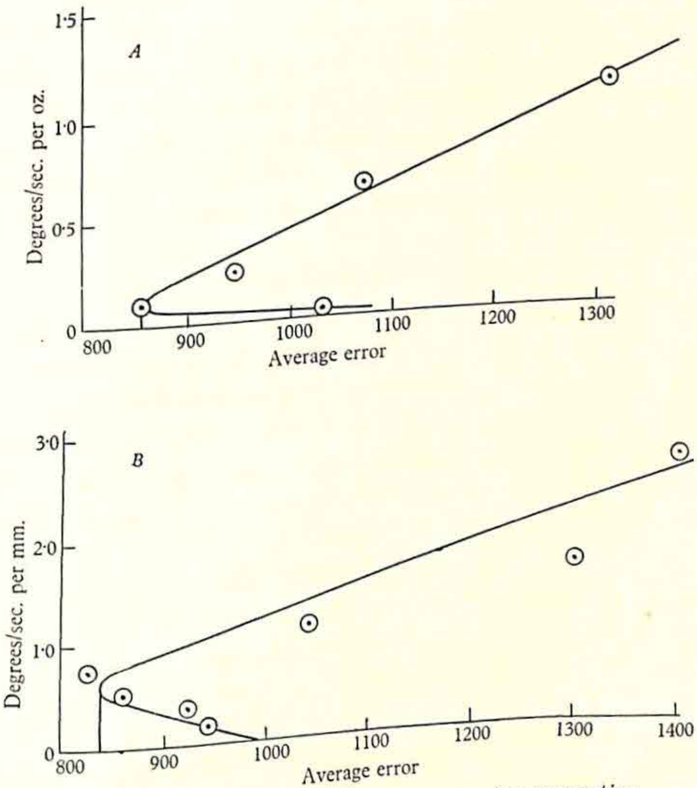


Fig. 4. A, pressure-gear ratios; B, free-moving-gear ratios.

the case of group B there was no deterioration in the scores after transfer from the free-moving to the easier pressure control. There was some disturbance due to the change of task, since the first score of group B after transfer to the pressure control was not equal to the final score of group A, who had learned this control as an initial task. But Fig. 3 shows that this disturbance was partly compensated by the relative ease[1] of the second task. Hence the first score of group B on the pressure control was not inferior to their final score on the free-moving control which they had practised directly. There was appreciable positive transfer from one control to the other in both groups, although the gross external responses were very different at the two controls.

The superiority of the pressure control was maintained after the interchange of groups and tasks, and errors were in the ratio of 3 : 4 in favour of the pressure control on the 6th day of practice. The difference was significant at $P < 0{\cdot}05$.

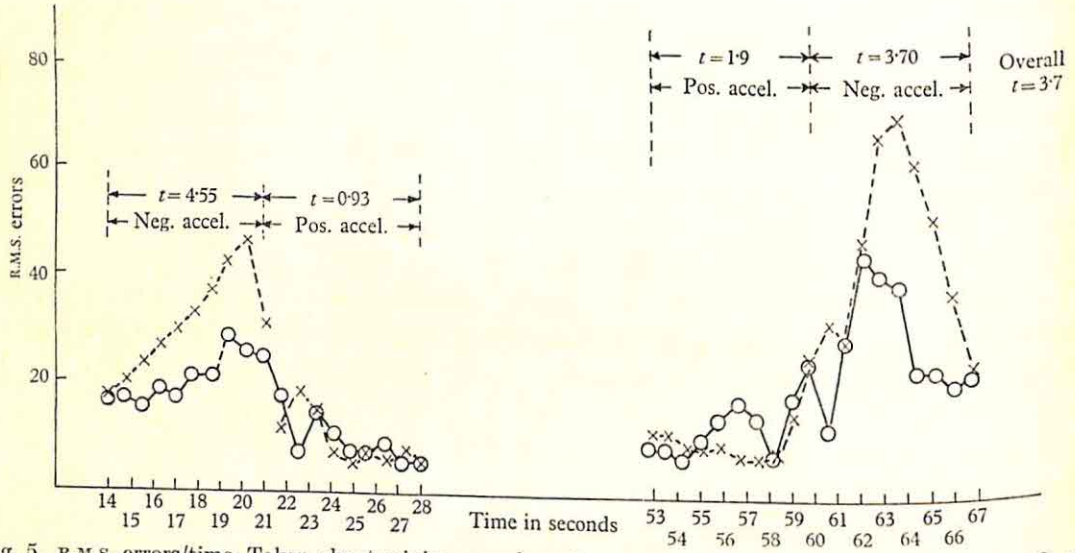


Fig. 5. R.M.S. errors/time. Taken about minimum and maximum rates of elevation course of Fig. 1. O—O—O, pressure, mean of five records; ×---×---×, free-moving, mean of five records. Records from five highly skilled subjects.

Comparisons after longer training. Twenty-four subjects, after longer preliminary practice, then did a further series of trials to compare the two levers on the difficult course of Fig. 1. These tests established the advantage of the pressure control at a high level of probability, $P < 0{\cdot}001$.

Graphic records of the elevation errors were taken from these subjects at the same time as the counter readings, to make sure that the method of scoring was not determining the answer being obtained. The graphic records from five of the subjects were analysed in detail, and Fig. 5 shows that the pressure control was again superior using the root mean values of the squared errors as the criterion. Although the records were taken from different people, they had striking similarities and showed clearly that the pressure control appeared to greatest advantage when it was relaxed towards centre under the influence of spring loading (i.e. during negative target accelerations) and also at cross-over points in the target course. Large tracking errors developed at these cross-over points since there was a sudden change in the direction of target motion, requiring a corresponding change in the continuous tracking adjustments of the limb.

---

[1] Definitions of the difficult-easy continuum in skill have previously been given (Gibbs, 1951).

During the 7 sec. of negative acceleration (towards the minimum velocity of the 1·8°/sec. of Fig. 1) the R.M.S. errors for the pressure control were compared with those for the free-moving control and a 't' value of 4·5 was found ($P<0·01$). But during the next 7 sec. of positive acceleration away from this speed there was no significant difference between the controls. There were similar effects about the maximum required speed of the course (i.e. 9°/sec. in Fig. 1). Again there was no reliable difference between the controls during the 7 sec. of positive acceleration, but the 7 sec. of negative acceleration following the peak speed showed a great advantage in favour of the pressure control ($t=3·7$, $P<0·01$).

Fig. 5 shows that the operators started to diminish the major errors at the cross-over points about 1 sec. sooner with the pressure than with the free-moving control since the peaks of the two error curves have a temporal separation of about 1 sec. at the cross-over points.

(c) *Finger operated controls*

A further group of fourteen subjects were used to compare the controls when operated by the fingers after equal practice in pressure and free-moving conditions.

The pressure control was then markedly better than the free-moving control with the easy course set at the apparatus ($t=4·6$, $P<0·01$). The average scores were slightly better when the controls were operated by the fingers as compared with the forearms, but the difference was not statistically reliable, since $t=0·73$.

The finger-operated pressure and free-moving controls were again compared, with the difficult course set on the apparatus. The superiority of pressure control was again highly significant at $t=5·85$, and $P<0·01$.

(d) *The single-correction series experiments*

Sixteen subjects compared the pressure and moving controls when large discrete corrections were made to a high level of accuracy. There are analysis difficulties in reaction and movement times, and a large number of recordings were therefore made, interpreted, and cross-checked by two observers. These tests showed that significantly less time was required to operate the pressure lever than the free-moving lever, after the signal to commence a correction ($t=3·44$, $P<0·01$). The time taken to *complete* this correction to a given level of accuracy was also significantly less when the pressure control was used ($t=6·9$, $P<0·01$). There was agreement between the 1 sec. difference in completing an accurate single correction and the 1 sec. displacement in the peaks of the two tracking curves of Fig. 5. Major factors in pressure superiority were therefore the reduced delay in modifying a previous muscular state, and settling more rapidly to a new rate of muscular contraction with this control. The results are set out in Table 1.

Table 1.

Condition	No. of subjects	No. of trials	Average time to start (sec.)	95 % confidence interval[1]	Average time to complete (sec.)	95 % confidence interval
Pressure	16	64	0·313	0·047 to 0·27 sec.	3·073	0·74 to 1·34 sec.
Free moving	16	64	0·426		4·113	

---

[1] The average difference was 0·113 sec., but this value is subject to sampling errors, whereas there is 95% confidence that the real difference lies between 0·047 and 0·27 sec.; similarly, the real difference in completing a correction was 0·74–1·34 sec. The times are rather long compared with some other tests of reaction and movement times, but this is due to the long time-constant of the apparatus, the ambiguity of the display, and the high level of final accuracy which was demanded.

VI. THE TABETIC PATIENT

A tabetic patient, aged 55 years, was tested on the apparatus. He was normally right-handed and usually employed as a general labourer. At the time of the tests the general reflexes were present but diminished, and the plantar response was negative. Sensitivity to pin prick was diminished at both hands and wrists, especially on the left side. The proprioceptors were impaired at the ankles and knees as shown by swaying on the Romberg test, and by diminished appreciation of movement and passive position. The pupils were typical of tabes dorsalis, and there were positive serological findings in the blood.

There was some weakness of the extensor muscle of the left wrist and wastage of the left hand. The patient could not adequately regulate the grip of this hand; but it could be used when sweeping with a broom, and for supporting weights such as a bucket of coal.

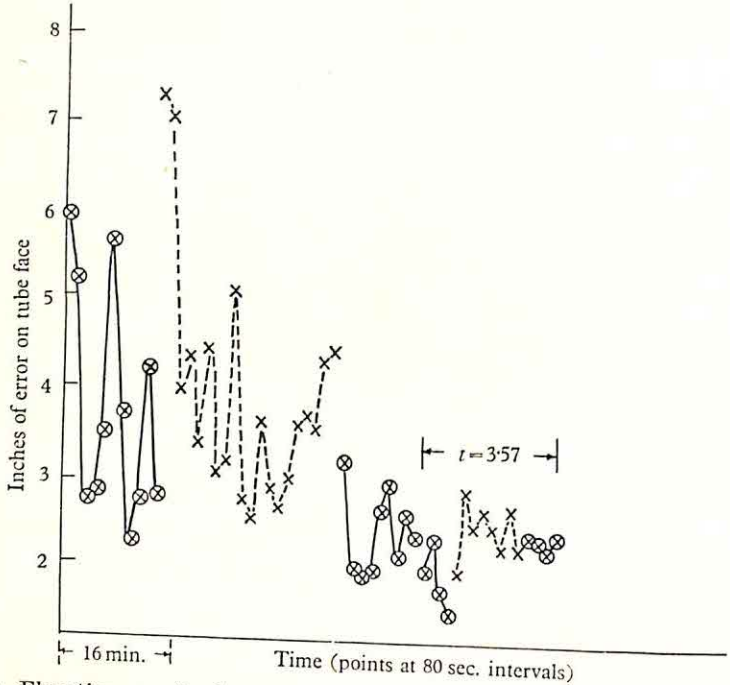


Fig. 6. Errors/time. Elevation record only on counters learning curve of tabetic patient. Fingers on right hand (intact). (Plotted points show total errors per 80 sec. practice.) ⊗—⊗—⊗, pressure, right hand (intact); ×- - -×- - -×, free-moving, right hand (intact).

The preliminary clinical tests of his ability to discriminate with eyes closed between different directions, rates and ranges of limb movement showed some impairment of proprioceptive sensibility in the finger joints of the right hand. This was slight, however, as compared with fingers of the left hand where there was considerable loss of sensitivity. Impairment was progressively less marked at the elbow and shoulder joints of the left arm.

The subject first practised tracking, using only the fingers and wrist of the intact right hand. The total error at the cathode-ray tube was 6 in. during the first 80 sec., but this was reduced to about 2 in. per 80 sec. at the end of 5 days of continued practice (Fig. 6).

The control was then transferred to the fingers and wrist of the impaired left hand, and there was an immediate and marked deterioration in performance with 13 in. of error at the tube face during the first 80 sec. of practice after transfer (Fig. 7). This was more than twice the initial error with the intact hand, despite the usual finding of considerable bilateral transfer in this kind of skill (Woodworth, 1938).

Seven minutes of practice were given in these conditions, and the placement of the control was again altered so that the fore and upper left arm could now operate the lever.

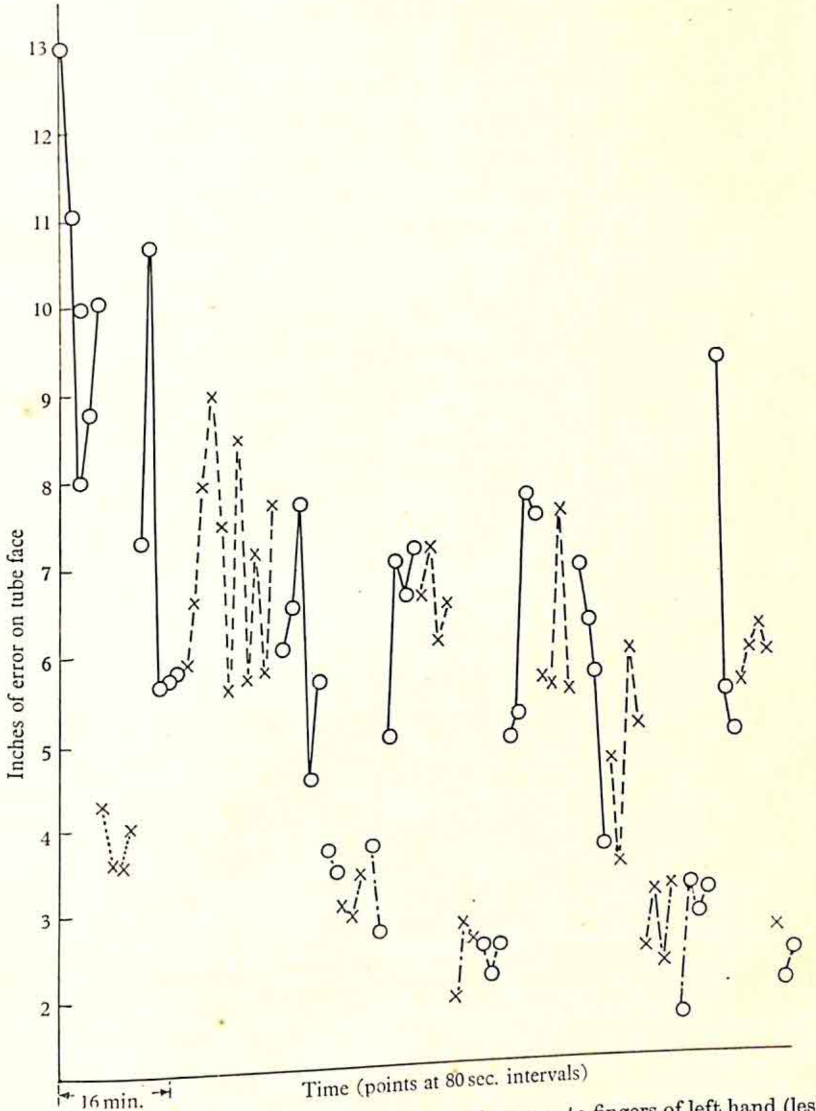


Fig. 7. Errors/time. Elevation record only on counters. Transfer curve to fingers of left hand (lesion) of tabetic patient, with check readings of fingers of right hand (intact) and four readings of transfer to whole left arm (intermediate impairment) (plotted points show total errors per 80 sec. practice). O—O—O, pressure, left hand (lesion); ×-- ×-- ×, free-moving, left hand (lesion); ×--- ×--- ×, free-moving, whole left arm (intermediate impairment); O—·—O—·—O, pressure, right hand (intact); ×—·—×—·—×, free-moving, right hand (intact).

The clinical tests had indicated intermediate impairment at these joints, and this was confirmed by the tracking scores, since the errors fell to 4 in. of error per 80 sec. after this second transfer. Five minutes' practice was given in the upper arm conditions, and the control was again reset so that practice continued with the impaired left fingers and wrist for a further 6 days. There was some improvement over the first 45 min. of practice with the impaired hand, and the subject learned to make periodic corrections in the correct sense so that the errors were reduced to about 5½ in. per 80 sec. But the rates of muscular contraction were seldom appropriate. In general, there was abnormal latency in initiating large corrective adjustments, and slow and continuous oscillations of the spot of light about the correct aiming point, a condition of instability which is also characteristic of servo mechanisms in certain conditions of phase lag in their feed back.

Only one tabetic patient was available for the tests, but both the intact and impaired states were found in one individual, thus equating the usual variables of intelligence, motivation, learning, etc.

The main purpose of these tests was to confirm the transfer results which had been found with intact subjects, when the level of proprioceptive stimulation was altered by mechanical means. For the rest, the results confirmed that vision can compensate only to a limited extent for the defects of locomotor ataxia; and the results are consistent with the findings of Sherrington (cf. Fulton, 1949), and also with the demonstrations by Holmes (1922) of disturbances in the direction, range, rate and force of movement following unilateral cerebellar lesion. Such agreement would be expected because of the heavy proprioceptive input to the cerebellum.

VII. Discussion

Pressure superiority is not explained by motor effects at spinal level

The considerable difference in the proprioceptive discharge during isotonic and isometric contraction could reasonably be regarded as the main experimental variable. There was more rapid release and settling to new patterns of contraction with the pressure control (Fig. 5; Table 1) to suggest that the discharge had important central effects, in addition to any smoothing and inhibitory effects at spinal level.

It was, however, important to consider first the possible advantages of pressure control in the motor aspects of performance. The movement range was negligible (0·08 in. maximum), and the time required to make such movements might be very small. This possible advantage could be discounted, however, in a paced task such as tracking where the correction speeds are externally imposed and little is gained by providing a control which can potentially exceed the target schedule. The sensitivity tests (Fig. 4B) showed that tracking errors were in fact increased by reducing the movement range below an optimal value, despite the assumed reduction in movement time. A lever with a small movement range could only be operated effectively when pressure cues were provided to compensate for the lack of movement cues (Fig. 4A, B).

Again, the errors of overshooting or leading a moving target might conceivably be reduced by a spring-loaded lever which opposed the dominant direction of tracking movements. But Fig. 5 shows that the pressure control was superior when spring loading was acting to accelerate, rather than to retard, the limb. Spring loading would tend to alter the force distribution in the muscles (i.e. controlled relaxations might tend to replace contractions in some phases), but the required force change was negligible in either case as compared with the maximum force output which can be developed if necessary in about 0·04 sec. Any time difference between generating and dissipating such force output is not of an order to explain the differences which were in fact found between the controls.

Motor effects, or sensory effects at spinal level therefore do not explain any significant proportion of the results, whereas a complete explanation is possible, if it is postulated that the sensory discharge is used continuously at central level, to regulate the force output in muscles.

Current hypothesis of the impulse discharge of movement

It is by no means generally believed, however, that the kinaesthetic discharge has important central effects, and the hypothesis of impulse discharge is currently supported by Lashley (1951) and Chernikoff & Taylor (1952). Their main objections to continuous feed back are based upon the somewhat conflicting evidence concerning kinaesthetic reaction time: about 0·125 sec. may elapse between an unexpected kinaesthetic stimulus and the release of an appropriate voluntary response (Craik, 1947; Vince, 1948; Hick, 1949; Chernikoff & Taylor, 1952). Lashley (1951) discusses the real difficulty of reconciling that latency with the observed times of movement, e.g. a single finger may be tapped at the rate of 10 strokes/sec., and the successive finger movement of a skilled pianist can be made in definite order at about 16 strokes a second. He concludes that 'the sensory control of movement seems to be ruled out in such acts'. The higher motor centres must release unmonitored impulses which have no effective sensory correlates if feed back lags are indeed constant at 0·125 sec.

Objections to the impulse hypothesis

The initial assumption might well be challenged, since the reaction times to unexpected stimuli can scarcely be transposed to practised skills or to the monitoring[1] of a movement which has already been released, but impulse theory may best be discounted on more general grounds. The inevitable implications of the theory have been clearly stated by Chernikoff & Taylor (1952): 'the control of movement consists of the perception of error, selecting a response pattern judged on the basis of previous experience to be appropriate to null out the error, and triggering the entire sequence'. But a rapid movement has a continuous phase, which is certainly not monitored by visual cues (cf. Fitts, 1951) and the lag in obtaining visual knowledge of results is particularly marked in such skills as golf, bowling, etc. If kinaesthesis is also to be excluded, then the proposal of impulse theory is that an appropriate response can be learned, and selected at will from a repertoire, without any sensory discrimination of its suitability until the initial phase of movement is completed.

The theory of knowledge of results is essentially a theory of feed back and is not compatible with the notion of impulse discharge. Observations of high-grade skill show that the feed-back stimuli which are learned as a basis for control must necessarily precede the visual knowledge of error. The feed-back stimuli of visual error tend to disappear as the correct anticipatory cues are learned, and this common observation points to two relatively distinct sources of feed back. It is no purpose of skill to acquire knowledge of visual errors; it is, rather, to find the means to eliminate such errors. It is therefore significant that the spino-cerebellar tracts conduct at the extremely high rates of 140 m. a second, so that only a few milliseconds are taken for conduction from dorsal roots to cerebellum.

Centripetal conduction time is probably significant in the learning of high-grade skills, but the time taken to act on sense data is much longer and is the critical time in the actual control of movement. Feed-back theory can give a reasonable basis for learning, but it must also suggest a means whereby the kinaesthetic discharge could be used to monitor and terminate the initial phase of a released movement, despite the known transmission lags of the central nervous system.

---

[1] The term monitoring is more specific than feed back. In this context it is defined as a process of continuous comparison between an intended state and the current state: continuous corrections are based upon the discrepancy and act in the sense which is appropriate to reduce the error.

The characteristics of proprioceptive nerve endings

The characteristics of several receptors have been determined by Matthews (1933) and must be considered in developing a theory of the control of movement. In resting or static conditions, the frequency of discharge in the muscle receptors is proportional to the logarithm of tension on the muscle. However, when the muscle is stretched, that simple relationship is disturbed and there is then only a correlation between the rate of change of frequency and the rate of stretching the nerve ending, but not between the absolute frequency and the absolute amount of stretch at any instant. The peak discharge during stretch, within the limits of the rates which were tested, may be over four times the final resting discharge in primary nerve endings, and more than twice as large in the tendon organs which lie in series with the muscle. Boyd (1953) has demonstrated a similar effect in the discharge from proprioceptors in the knee joint of a cat. The known discharge in isotonic conditions can therefore only define the rates of movement during a rapid movement; the final resting state of position is only defined by the discharge after movement has ended, when the receptors have again settled to equilibrium.

There is no rate of stretch component to complicate the discharge during isometric contraction, and the discharge therefore continuously defines the important control quantity of present state, since the absolute frequency is logarithmically proportional to the absolute applied tension. This discharge could define both the present tension, and rate of change of tension in the muscles.

Lissman (1950) quotes data by Katz, that the discharge in some primary endings is considerably boosted in isometric conditions. There is reason to believe that the isometric discharge is carried in fibres of large diameter whereas the secondary nerve endings, with a low threshold to rate of stretch, have smaller fibres and are inactive during isometric conditions. Isometric sense data may therefore be conducted more rapidly, and may be more adequate in control because of the additional degree of information in this discharge.

The action of cutaneous receptors is not important in the present context, since they have little effect in the control of movement (Sherrington, see Fulton, 1949). Subsidiary experiments by the present writer have shown that the accuracy of tracking with pressure control is not impaired when the controlling fingers are strapped very tightly to the control lever.

A hypothesis of feed-back control

It is perhaps axiomatic that movements of definite extent cannot be made without knowledge of the starting and end position of the limb. Vision or proprioception could adequately define the initial position of a stationary limb, but the quoted evidence (Fitts, 1951) shows that only proprioception need be considered as a possible means of first halting a rapid movement by current sensory control.

A rapid movement cannot be terminated by a system of kinaesthetic 'position' indicators which might conceivably be arranged to indicate that a desired position or angle had been reached. Such movements would be overshot because of the time taken by feed-back impulses to traverse the higher centres and to return as motor impulses (estimated by Ruch, 1951, as 19 to 36 msec.) plus the time taken by contractions to arrest the moving limb. Positional control is more definitely precluded by the action of the known receptors. A more reasonable principle of current control is that a rapid movement is first halted by integrating, over a period of time, the feed-back signals of the rate of movement, i.e. the extent of movement could be estimated by integrating a known rate of movement over a known period of time.

The principle of the rate control of movement would imply that in the apparently continuous task of tracking, there are only intermittent estimates of the successive positions reached by a moving limb. Such tasks would therefore have definite temporal phases, corresponding to the length of the integration period, and the end of one phase of movement is necessarily the starting-point of the next. There would be a definite limitation in such a system, that a released movement must necessarily run for an integration period to give the initial data for the next phase, since a new movement of definite extent cannot be organized without knowledge of its starting-point.

The hypothesis would completely account for the present results, that, with the pressure control, there was more rapid response to a sudden large error and quicker settling to a change of target behaviour. This was probably because the isometric discharge continuously defines the 'present' state of tension in the limb. The delays which are necessarily involved in any process of integration are presumably reduced by the extra degree of information in the isometric discharge. The extensive general evidence of control intermittency has been reviewed by Welford (1952). It is suggested that the present results can best be regarded as a special case of the psychological refractory phase, and the present hypothesis may therefore account, at least in part, for this general effect in movement.

The principle of rate control clearly disposes of the main objections to feed-back theory, i.e. that the known transmission latencies of the central nervous system preclude the possibility of current sensory control. The duration of movement (i.e. the integration period) is predetermined to allow for these latencies, and the continuity and rate correlation of the feed-back impulses permits purposive control and monitoring of the rates of movement before the limb has reached its final and intended position.

The hypothesis may also throw light on the reasons for the well-known tendency toward time constancy when making movements of very different extent (cf. Fitts, 1951). The two controlled variables of a velocity system are the rate and the duration of movement. Clearly, the difficulties of learning a repertoire of movements of different extent would be greatly reduced by a tendency to keep one controlled quantity constant and learning only a range of variation in the other. Hence time constancy may well appear as a device for greatly reducing the repertoire of learned items, while still retaining the required versatility.

Craik (1947) has previously suggested that anticipation is essentially a process of extrapolation from current sense data. It is clearly impossible to extrapolate from a present to a future bodily state without current knowledge of the rate of change of state. Hence, the continuity and rate correlation of the kinaesthetic discharge gives a continuous sensory basis for anticipation, as would be required by Craik's hypothesis.

Learning and transfer effects

The hypothesis of velocity control is therefore entirely consistent with physiological findings, the dominant characteristics of skilled movement, and with the theory of knowledge of results. But the consistency of feed-back theory does not necessarily preclude the alternative notion of ballistic impulse, and, in this connexion, the results of transfer experiments are of decisive importance.

The present experiments were undertaken largely to confirm and explain previous data on the transfer of skill (Gibbs, 1951) which had shown that physical similarity of stimuli and responses in a first and second task was not a particularly meaningful element in the transfer of high-grade skill. Of greater importance were the optimal zones in a range of exteroceptive or interoceptive sense data (e.g. Fig. 4A, B), and the relative adequacy of the sense data in the first and second task was an important determinant of the amount of transfer between them. These earlier findings are confirmed by the results of the present experiments since there was no deterioration in group performance after transfer from the free-moving to the pressure control (i.e. from meagre toward adequate sense data), whereas there was a sharp decline in performance after transfer with the opposite order of presentation. Again there was little transfer of skill from the intact to the impaired hand of the tabetic patient (Figs. 6 and 7). The previous experiments also showed high positive transfer between two different tracking courses, which presented entirely different problems in space and time, indeed, the relative content or difficulty of such tasks could be arranged so that indirect was equal to direct practice.

Such results show clearly that the organism is not 'calibrated' by learning. Stimuli and responses are not learned as a record of events in space and time, nor do specific external stimuli trigger a blind series of stereotyped muscular contractions, which run their course independently of current sensory control. Skilled behaviour can be rapidly adapted to different conditions in space and time, but the greatest transfer is obtained when there are equally adequate kinaesthetic data in the second task to define the changed conditions at the periphery. These data are fed back continuously and define the departure from a desired condition, so that the organism is in fact continuously actuated by error during a skilled movement.

It seems, therefore, that a main purpose of learning a skill is to filter and to interpret the anticipatory cues of the kinaesthetic discharge. The periodic oscillatory errors in the early stages of learning a tracking task (e.g. riding a bicycle) are well known, and are probably due to dependence on visual feed back, with consequent delays in knowledge of results. Oscillatory errors were very marked in the records of the tabetic patient throughout his training, and the stability of skilled performance can easily be disturbed by the use of high sensitivity controls (Fig. 4A, B), which have inadequate movement or pressure cues.

The problems of learning, in relation to high-grade skill, may therefore hinge upon relatively small time differences in the feed back from vision and kinaesthesis, and upon the anticipatory characteristics of the kinaesthetic discharge. The learning and sensitivity experiments, and the sluggish responses of the tabetic patient do not support Lashley's (1951) suggestion, that the rapidity of skilled movements precludes the possibility of current sensory control. It seems much more probable that the exact converse is true; the need for knowledge of results from kinaesthesis sets a real limit to the rates of successive innervation in synergic muscle groups.

VII. SUMMARY

Ninety-five normal subjects and one tabetic patient were tested in experiments which were designed to alter systematically the response in a visuo motor skill and, thereby, to alter the proprioceptive discharge in the controlling muscles.

Two types of joystick were used to control the output of a velocity control servo mechanism in a tracking task. One lever was of the normal free-moving type, and there the muscular contractions were isotonic when the subject was making controlled responses by displacing the lever. The other had strong spring centring, but it could be deflected slightly by applying pressure, so that the controlled contractions of muscle were virtually isometric. Changes of limb position, and changes in the force exerted by the limb could therefore be compared, using the criteria of speed and accuracy.

Isometric contraction with the pressure control was definitely better when making both discrete and continuous responses, learning was easier, and differential transfer effects appeared which also favoured the pressure control. It is thought that this superiority is due to the marked difference in the proprioceptive discharge in the two contractile states. The immediate implications are that skilled movements are continuously regulated by using kinaesthetic data which are generated by the limb in motion and are continuously fed back to the higher motor centres. The broader implications are briefly discussed, especially how cause and effect operate in sensory organization: a theory of control is therefore developed.

My thanks are due to Dr W. E. Hick and Dr N. H. Mackworth for reading and discussing the paper, and to Mr N. L. Webb for advice on the statistical analysis. Dr D. Russell Davis kindly made available the help of the tabetic patient and his clinical data.

References

Bell, C. (1826). On the nervous circle which connects the voluntary muscles with the brain. *Phil. Trans.* pt. 2, 163–75.

Boyd, I. A. (1953). Nerve impulses from proprioceptors in the knee joint of a cat. *J. Physiol.* cxix, 8/*P*.

Chernikoff, R. & Taylor, F. V. (1952). Reaction time to kinaesthetic stimulation resulting from sudden arm displacements. *J. Exp. Psychol.* xliii, 1–8.

Craik, K. J. W. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.* xxxviii, 56–61; 1948, xxxviii, 142–8.

Fitts, P. M. (1951). Engineering Psychology and Equipment Design. A chapter in the *Handbook of Experimental Psychology*, pp. 1328–9. Ed. S. S. Stevens. New York: J. Wiley and Sons.

Fulton, J. F. (1949). *Physiology of the Nervous System*, pp. 168 and 518. New York and Oxford.

Gibbs, C. B. (1951). Transfer of training and skill assumptions in tracking. *Quart. J. Exp. Psychol.* iii, 99–110.

Hick, W. E. (1949). Reaction time for the amendment of a response. *Quart. J. Exp. Psychol.* i, 175.

Holmes, G. (1922). The Croonian lectures on the clinical symptoms of cerebellar disease and their interpretation. *Lancet*, 202, 1177–82; 1231–7.

Lashley, K. S. (1951). The problem of serial order in behaviour. A chapter in *Cerebral Mechanisms in Behaviour*, ed. L. A. Jeffries. New York: J. Wiley and Sons.

Lissman, H. W. (1950). Proprioception. A chapter in *Physiological Mechanisms of Animal Behaviour.* Soc. Exp. Biol. Symp. no. iv. Cambridge University Press.

Matthews, B. H. C. (1933). Nerve endings in mammalian muscle. *J. Physiol.* lxxviii, 1–53.

Ruch, T. C. (1951). Motor Systems. A chapter in the *Handbook of Experimental Psychology*, pp. 199 and 204. Ed. S. S. Stevens. New York: J. Wiley and Sons.

Vince, M. A. (1948). Corrective movements in a pursuit task. *Quart. J. Exp. Psychol.* i, 85–103.

Welford, A. T. (1952). The psychological refractory period and the timing of high-speed performance. *Brit. J. Psychol.* xliii, 2–19.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychol. Rev. Monogr. Suppl.* iii, 13.

Woodworth, R. S. (1938). *Experimental Psychology*, pp. 181–7. London: Methuen and Co.

(*Manuscript received* 15 *April* 1953)

THE ADEQUACY OF EVERYDAY QUANTITATIVE EXPRESSIONS AS MEASUREMENTS OF QUALITIES

By D. SHEPPARD

National Institute for Research in Dairying, University of Reading

Quantitative terms, such as 'Very' or 'Rather', are frequently used for describing qualities, both in everyday conversation, and by psychologists for certain types of rating scales. Admittedly such quantitative terms may be imprecise in meaning, yet it is generally considered that there will be fair agreement in the interpretation of these terms, provided that the quality which is being described is well known.

Psychologists have shown that judges may disagree markedly when they have to assess independently the qualities of any one thing, but it is not easy to tell from such evidence whether the judges disagree about the merits of the thing being judged, or whether they merely have different opinions about the meanings of the words with which they choose to describe it. In addition, it is not easy to find qualities that large numbers of people can judge expertly, and for which certain people will not have special preferences or biases, as for example with assessments on flavour.

For certain purposes, however, it is necessary to obtain absolute judgements for qualities of materials in terms of such quantitative descriptions. Thus when obtaining consumer ratings for food qualities at the National Institute for Research in Dairying, it was an easier task for the raters to use everyday quantitative terms to describe the values, than to allocate marks for these qualities by a points system.

A set of descriptions so obtained may be sufficient for some purposes, but in most cases it is desirable, if theoretically unjustifiable, to be able to allot numerical values to these descriptions, so that at least some simple statistical treatments can be used on them. The most usual procedure seems to be to allot serial numbers to the quantitative descriptions, and then to treat them as if they were equally spaced one from another, and this, although an arbitrary procedure, often leads to fruitful results.

In view of these considerations, some simple experiments were carried out in an attempt to find out how people agreed about the use of quality descriptions for items of everyday experience, and also to attempt to discover a better method of allotting numbers to these quality descriptions, for the purposes of statistical treatment.

METHOD

For very many reasons, it is hard to compare people's subjective scales by asking them how they would describe given samples of material. Thus, for example, there are known effects which make such a comparison difficult. Amongst these are the 'central tendency' (Hollingworth, 1910), the fact that ratings tend to be a function of the stimulus range presented (Tresselt, 1948; Sheppard, 1953) and other difficulties that are incurred for the reason that objectively equal intervals are unlikely to be subjectively equal.

An unusual method of approach, which avoids these particular difficulties, is to present the subject with a list of quality descriptions, and ask him to say what objective measure of the quality concerned corresponds to them. This can of course only be done when the subjects are familiar with the objective measurement concerned and are used to thinking in terms of it, and provided that the qualities 'form a continuum'.

For this purpose then, a number of people were asked the following questions:

Test A. How far is (*a*) very far, (*b*) rather far, (*c*) very near, (*d*) near, (*e*) far, (*f*) rather near? (*g*) What distance lies in between, and is neither far nor near?

These questions were asked in this order, and the experimenter noted down the answers.

Test B. If you were going for a walk somewhere, what distance would you regard as being 'very far', etc. Please fill in what you consider to be suitable distances into this rating scale.

Here the descriptive qualities were the same as in test A except that 'not very near' and 'not very far' were added. The 'positive' values were written out in a definite order against a horizontal line. The subjects read the descriptions and filled in the appropriate numerical values themselves in this test, and also in tests C and D.

Test C. If you were waiting for a bus, what sort of time would you have to wait for you to describe it as being 'a very long time', etc. Please fill in the appropriate times into this rating scale.

Here the method was the same as for test B, with the qualities as follows: (*a*) a very long time, (*b*) a long time, (*c*) a rather long time, (*d*) neither a long nor a short time, (*e*) a rather short time, (*f*) a short time, (*g*) a very short time, (*h*) a not very long time, (*i*) a not very short time.

Test D. How much would a single man have to earn, in a week, for you to describe him as being 'very rich', etc. Please fill in the appropriate sums into this rating scale.

Here the qualities were similar to those in test C, but ranging from 'very poor' to 'very rich'.

Test E. In addition, some 300 subjects were asked to fill in descriptive values onto a 7 in. straight line, the written instructions for this being as follows:

You must often have made remarks such as 'that is very good' when describing the quality of different things. Do you consider that, for example, 'good' is as much better than 'average' as it is worse than 'very good'?

A quality scale ranging from a point representing 'extremely good' at one end, to a point representing 'extremely bad' at the other, is given below. Suppose that equal distances along the line represent equal increases, or decreases, in quality, can you indicate where *you* think the following quality judgements should occur in this scale? Arrange the quality descriptions so that they are not only in what you think is the right *order*, but so that the *distances between* each quality description are what you consider to be correct.

Very good, Rather good, Good.
Very bad, Bad, Rather bad.
Average.
Not very good, Not very bad.

Use a ruler if you wish, and mark in the appropriate positions for each of these qualities by a stroke crossing the line, and name the intersection.

Please do not discuss this with anyone else, until you (and they) have completed the task.

Note. (1) The same point on the scale may be used for more than one quality description if you wish. (2) Certain of these descriptive terms may differ in meaning if one or other of their constituent words is emphasized. Try to consider them, therefore, as if no one word in each term were emphasized more than any other.

For certain of these forms the word 'Moderate' was substituted for 'Average' and the word 'Fairly' for 'Rather'. In other forms 'Quite', or 'Fairly', were put in place of 'Rather'.

A large number of subjects answered these questions, although different sized groups were used in each test. Thus only ten subjects did test A, eighty-five subjects did tests B, C and D, and approximately 300 subjects did test E. Test A was, in fact, only a preliminary one, and the largest number of subjects were used for test E since the results of this were of practical service to the writer.

Careful selection of subjects was not considered necessary for these tests, and the groups comprised service personnel, research workers, students and members of W.E.A. classes. Both men and women took part, and it seems likely that the groups were, on the whole, above average in intelligence. It is probable that less intelligent subjects would not have understood the tasks.

One group was carefully selected, however, for ninety-three teachers were asked to do test E. This was because they were thought to be 'expert' in the use of such qualitative terms, which are often used for school marking. The data provided by this group were examined separately. The group consisted of both men and women teachers from five primary, three secondary modern and two grammar schools.

RESULTS

These will be discussed in three sections, section 1 dealing with the results obtained from test A, section 2 with the results from tests B, C and D, and section 3 dealing with the results from test E.

Section 1

Only ten subjects were asked the questions in test A, and this test did little more than demonstrate how difficult subjects found such a problem, particularly when the situation was left completely ambiguous.

The answers varied enormously, 'Very far' being given as 100 miles by one person, and 1000 light-years by another, and 'Very near' being given as 200 yards by one person, and the diameter of a molecule by another.

There was one other interesting feature, in that the subjects often gave responses in an unexpected order, thus the distance for 'Rather far' would sometimes be greater than the distance for 'Far', and sometimes less.

Section 2

In tests B, C and D the situations were more carefully defined, but no more carefully than they would be in everyday conversation. Seventeen subjects who did these tests were presented with rating scales which were 'inconsistent' in that for one of the three tests the 'Rather' term on the scale was written in next to 'Very', instead of next to the term 'neither x nor y'.

From the results, it was obvious that most people of this group thought that 'Rather' should come next to 'neither x nor y', although only four out of seventeen subjects were careful enough to give this order consistently for all tests. Evidently, then, the word 'Rather' is an ambiguous one, and its interpretation will depend largely on the position which is given to it in the rating scale presented to the subjects.

Sixty-eight subjects were given tests B, C and D with the descriptive phrases in the more generally approved order, that is, with the 'Rather's' next to 'neither x nor y'.

A histogram for the results of test D, for these sixty-eight subjects, is shown in Fig. 1, in which the number of pounds, grouped logarithmically, are plotted against the frequency of responses for each of the descriptive terms. It can be seen that there is very considerable disagreement about the value for each of the terms.

There is one consideration to be taken into account, however, in that test D, in common with tests A, B and C, asks subjects to define a scale which is unlimited at one end, but limited at the other. Thus, for example, all subjects would agree that a sum of £1000 per week would make a single man 'Very rich', although many subjects gave less than this sum. In addition, there is naturally more scope for disagreement at the 'unlimited' end of the scale than the 'limited' end.

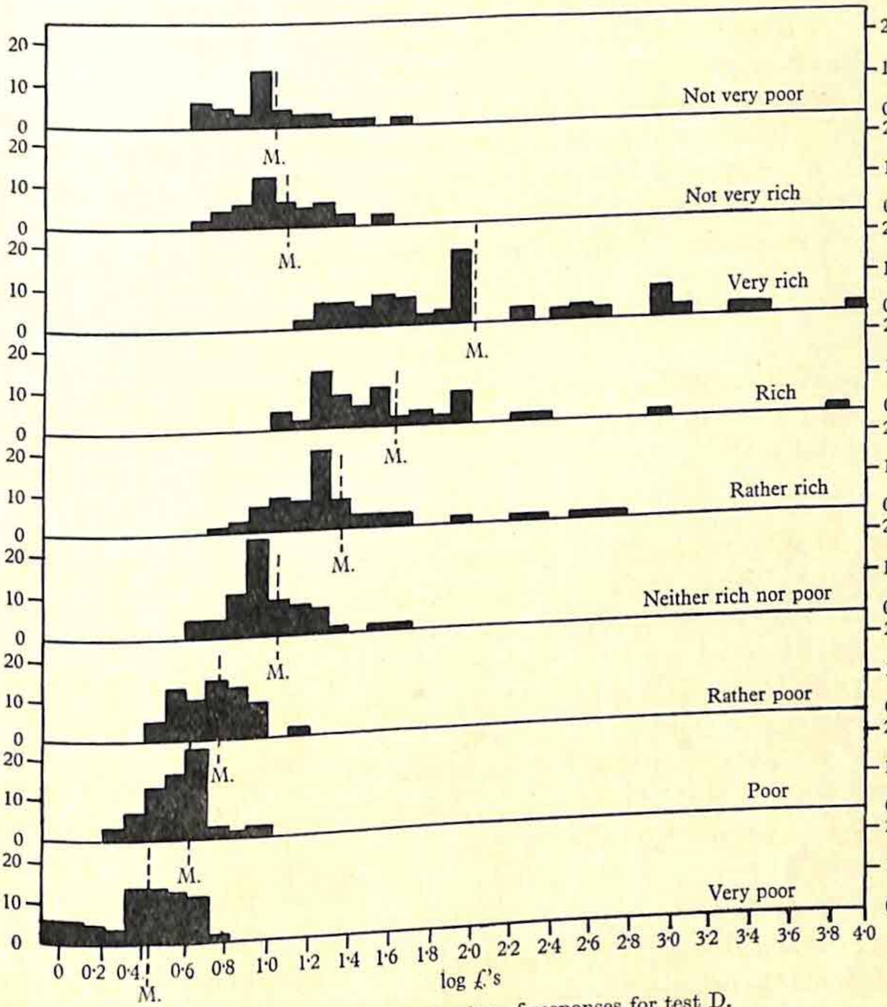


Fig. 1. Frequency distribution of responses for test D.

Table 1 shows the range of the responses for test A and the logarithmic means and standard deviations for tests B, C and D as done by the sixty-eight subjects. The extreme values which were given for these latter tests are also shown.

Plotting the data logarithmically was a reasonably successful procedure with the results given for tests B, C and D, but most successful with test D, where subjects tended to give values first of all in pounds, then in tens, in hundreds, and finally in thousands

Table 1. *The variations in the objectivity measurements given for qualitative values*

		Range of values given								
Test A	Descriptive term	Very near	Near	Rather near	Neither near nor far	Rather far	Far	Very far		
	Smallest value	4 × 10⁻⁹	1 yd.	¼ mile	150 yd.	6 miles	2 miles	100 miles	—	—
	Greatest value	200 yd.	2 miles	20 miles	63 miles	To the moon	To the sun	1,000 light-years	—	—
Test B	Descriptive term	Very near	Near	Rather near	Neither near nor far	Rather far	Far	Very far	Not very far	Not very near
	Smallest value	2 yd.	15 yd.	20 yd.	300 yd.	440 yd.	¾ mile	3 miles	300 yd.	200 yd.
	Greatest value	3 miles	5 miles	7 miles	10 miles	20 miles	50 miles	100 miles	11 miles	12 miles
Test C	Descriptive term	Very short	Short	Rather short	Neither short nor long	Rather long	Long	Very long	Not very long	Not very short
	Smallest value	1 sec.	30 sec.	1 min.	3 min.	5 min.	10 min.	15 min.	2½ min.	3 min.
	Greatest value	5 min.	10 min.	20 min.	30 min.	45 min.	2 hr.	3 hr.	30 min.	45 min.
Test D	Descriptive term	Very poor	Poor	Rather poor	Neither rich nor poor	Rather rich	Rich	Very rich	Not very rich	Not very poor
	Smallest value	8*s.*	£2	£3	£5	£7	£12	£16	£4. 10*s.*	£4. 10*s.*
	Greatest value	£5. 6*s.* 7*d.*	£10	£15	£50	£500	£4,000	£20,000	£40	£30
		Means and standard deviations for values grouped logarithmically								
Test B	Descriptive term	Very near	Near	Rather near	Neither near nor far	Rather far	Far	Very far	Not very far	Not very near
	Log mean (miles)	$\bar{1}$·11	$\bar{1}$·59	$\bar{1}$·97	0·32	0·64	0·87	1·10	0·16	0·30
	Mean	228 yd.	677 yd.	1640 yd.	2 miles 156 yd.	4 miles 625 yd.	7 miles 737 yd.	12 miles 985 yd.	1 mile 755 yd.	1 mile 1735 yd.
	Log S.D.	0·67	0·56	0·48	0·40	0·35	0·30	0·38	0·23	0·39
Test C	Descriptive term	Very short	Short	Rather short	Neither short nor long	Rather long	Long	Very long	Not very long	Not very short
	Log mean (min.)	1·89	2·37	2·71	2·88	3·15	3·38	3·61	2·92	3·03
	Mean	47 sec.	2 min. 21 sec.	5 min. 8 sec.	7 min. 35 sec.	14 min. 0 sec.	23 min. 59 sec.	40 min. 45 sec.	8 min. 55 sec.	10 min. 43 sec.
	Log S.D.	0·52	0·50	0·26	0·24	0·23	0·23	0·30	0·20	0·35
Test D	Descriptive term	Very poor	Poor	Rather poor	Neither rich nor poor	Rather rich	Rich	Very rich	Not very rich	Not very poor
	Log mean (£'s)	0·42	0·61	0·78	1·04	1·38	1·61	2·01	1·07	1·01
	Mean	£2. 12*s*, 4*d.*	£4. 16*s.* 7*d.*	£6. 5*s.* 2*d.*	£10. 19*s.* 1*d.*	£23. 15*s.* 5*d*	£40. 14*s.* 9*d.*	£101. 8*s.*	£11. 13*s.* 5*d.*	£10. 4*s.* 1*d.*
	Log S.D.	0·20	0·14	0·18	0·19	0·41	0·41	0·61	0·19	0·23

Test ...	A	B	C	D
Number of subjects	10	68	59	68

Table 2. *Means and standard deviations for 'abstract' scale* (*test E*)

Test E	Descriptive term	Very bad	Bad	Rather bad	Average	Rather good	Good	Very good	Not very good	Not very bad
	Mean in inches	6·3	5·5	4·9	3·4	2·3	1·7	0·8	3·4	3·9
	S.D. in inches	0·32	0·44	0·70	0·29	0·47	0·67	0·27	0·57	0·58

Number of subjects 78

of pounds. With test B, where the scale started off in yards, through fractions of miles, miles and tens of miles, this conversion was not so successful.

Evidently then, the fact that certain expressions come more readily to mind than others limits the value of data obtained from such an experiment. As an example, subjects all describe the earnings of a 'Rich' single man in multiples of £5.

Even despite this consideration, however, it remains true that the values given for the contiguous qualities are logarithmically related to one another, at least approximately, so that this might indicate that the Weber law, or something like it, can be applied to real-life mental schema.

As Wolters (1916) has suggested, the negative concepts proved to be interesting in these tests. Only forty out of the sixty-eight subjects inserted the negative phrases at all in test D, and of these eleven considered that to be 'Not very poor' was about the same as being 'Not very rich', twenty-three considered that to be 'Not very poor' implied that one was poorer than 'Not very rich' and six held the contrary opinion.

'Not very rich' was held to mean anything between 'Very rich' and 'Rather poor', with thirty-three subjects placing this between 'Rather rich' and 'Rather poor'; 'Not very poor' was held to mean anything between 'Rich' and 'Poor', but thirty-seven placed it between 'Rather rich' and 'Rather poor'.

When the results for different scales were compared for different individual subjects, however, it was found that the negative judgements were not consistently placed, either in their relation to one another, or in their relation to the positive judgments.

Section 3

Excluding the teachers, 118 subjects were given these particular phrases for test E; of these some 73% used the order V.G., G., R.G., Av., R.B., B., V.B. Histograms for the results are shown in Fig. 2; in which the position of the intersection on the 7 in. line is plotted against the frequency of occurrence.

The means and standard deviations for these data have been calculated since the distributions are in most cases approximately normal about a central point. These are given in Table 2. In some cases these may be misleading, however, since for certain qualities the distributions appear to be skewed.

No doubt some of the apparent disagreement results from subjects inserting their points inaccurately, but even with the term 'Average', for which there is more agreement than for any other, it is evident that some subjects cannot have intended that it should come at the mid-way position.

As before, there is less agreement about the placing of the negative terms, than for the positive ones.

Some 14% of these subjects gave the order V.G., R.G., G., Av., B., R.B., V.B., some 3% omitted certain of the terms since they did not feel sure of them, and some 10% were suggestible enough, or careless enough, to use the order given in the instructions, i.e. V.G., R.G., G., Av., R.B., B., V.B., an order that had been chosen so as not to influence the subjects' placing of 'Rather'.

Plotting the frequencies in a histogram form obscures the individual differences in the results. A large number of the individual results were therefore plotted out, as in Fig. 3; but it was not possible to find sufficient common features to be able to group them satisfactorily in any way. Some of the curves which resulted were linear and some were approximately sigmoid, or inverted sigmoid, in shape; but for the majority it was only possible to say that the 'Bad' side was a mirror image of the 'Good' side. Some were completely irregular, as for example scale (*d*) in the figure.

There were no apparent differences in the subjective scales when the terms 'Fairly' or 'Quite' were substituted for the word 'Rather', except that there was more agreement that the former terms should come next to 'Average' than there was for the latter. The completed scales for which 'Moderate' and 'Fairly' were included were also very similar to those given for the other terms.

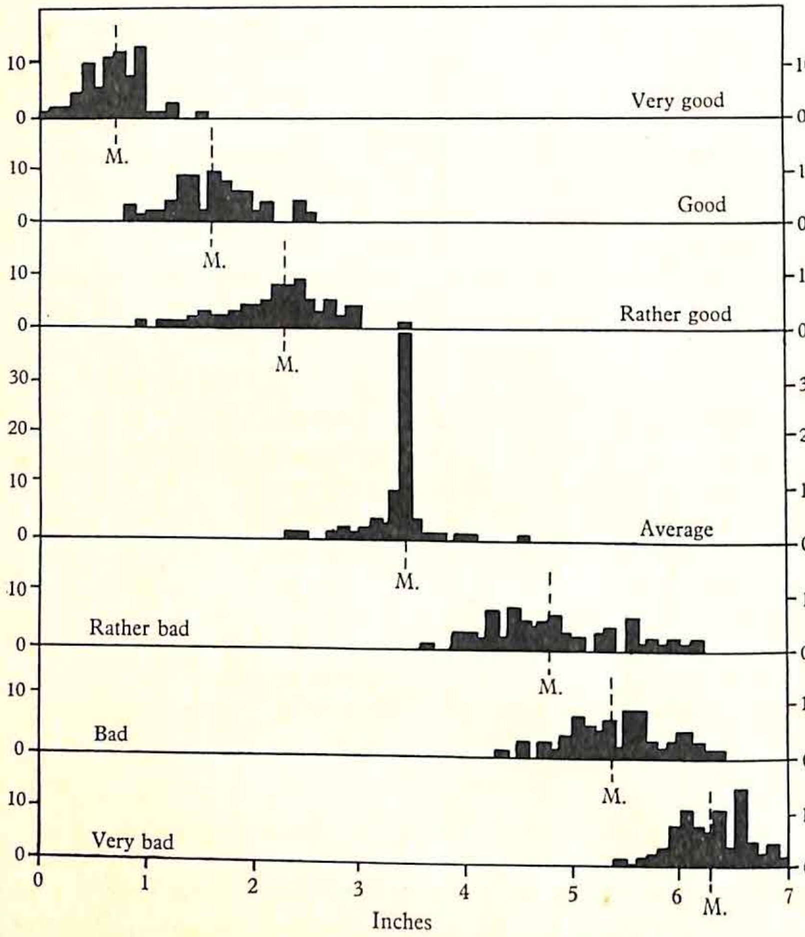
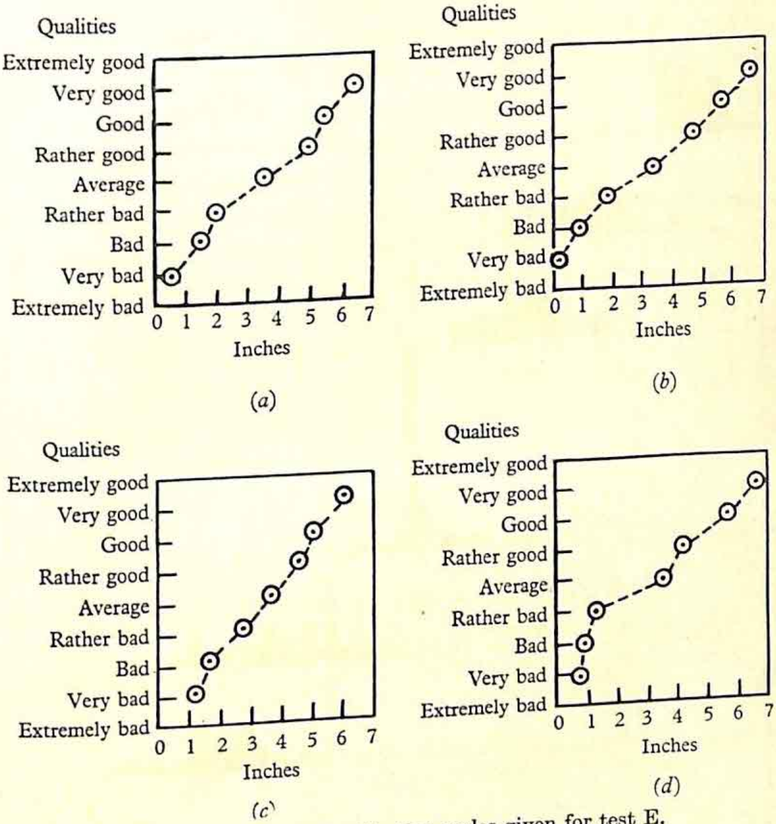


Fig. 2. Frequency distributions of responses for test E.

It is interesting, however, that although there was, in a sense, agreement between the groups, yet the resulting scales, when averages were plotted, were not linear ones.

Some subjects criticized these tests as being too abstract, but in fact they did seem to have some meaning, since it would have been possible to predict with some success the relation between the values given in tests B, C and D for different individuals, from their scales given by this test. Furthermore, it was possible to make a much smoother curve relating subjective judgements on dairy products to an equivalent objective measurement, by using a non-linear scale derived from tests of this type, than by treating the subjective values as though they were equally spaced from one another. In addition, the smoother curve was more like that which would have been expected if the Weber law was in operation (Prentice, 1953).

By way of interest, a large number of teachers were asked to do test E, since school marking is often done in such qualitative terms. Forty-three teachers were given the scale in which 'Rather good' and 'Rather bad' were included, but since they objected to the use of these terms, a further fifty teachers were asked to do the test with the term 'Fairly' substituted for 'Rather'. A few teachers even objected to this, and said they would use 'Fair' rather than 'Fairly bad', while nearly all objected to the use of the negative terms.



(a) (b) (c) (d)

Fig. 3. Examples of subjective scales given for test E.

The scales which resulted were, on the average, only a little more linear than those given by other groups, and there was no more agreement about the positions on the line for the terms, between teachers in different schools, or even in the same schools, than there had been with the other groups of subjects. In Fig. 4, three scales given by teachers in the same school are shown, and it is easy to see that a child would be less likely to get a good mark from teacher C than from teacher A.

It was interesting to see subjects' reactions to these problems; thus some took the task very seriously, and made several attempts before being satisfied; some at once thought of reasons why it could not be done, or why the results would be meaningless; and a small number refused to, or were quite unable to do it. The latter were mostly people who take themselves rather seriously.

A good many points of interest were raised in comments. Thus three subjects thought that 'Bad' should have been on the left, and 'Good' on the right; one subject thought that the descriptions were not all in the same dimension; some defined the qualities by marking in two points, the quality in question applying over this distance; and very many objections were raised to the negative terms.

Discussion

Evidently the results of these tests show that such quantitative terms for describing qualities are very ambiguous; yet for certain purposes, as for example with rating scales, it may be better to use such terms, which everyday people can understand, than to ask judges to attempt the more abstract task of judging by a points system.

A method such as that used in test E allows a somewhat less arbitrary interpretation of such quality descriptions and, as previously stated, it has been found to be useful in relating subjective assessments to objective measurements. Furthermore, it seems that by using the somewhat ambiguous terms 'Good' and 'Bad' one can build up a scale which is as limited in one direction as it is in the other; while it is also possible to obtain a relation between the quantitative terms which is sufficiently independent of the quality (in this case 'Good' or 'Bad') for it to be possible to transfer this relation and apply it successfully to assessments for another quality. Thus this was possible in the examples given here, for the subjective tests on food qualities, and for the qualities represented in tests B, C and D.

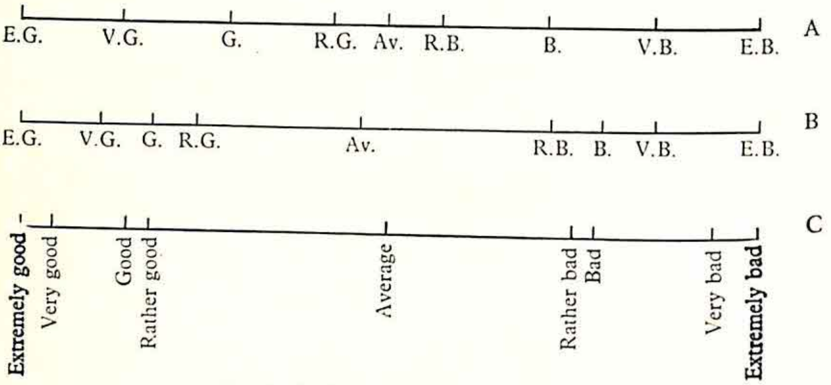


Fig. 4. Scales given by three teachers.

It can be seen from the histograms in Figs. 1 and 2 that if *complete* agreement is required between subjects, nothing but a three-point scale using the two 'Very' terms and 'Average' could give this. But a technique such as that for test E does at least make it possible to interpret the meaning of descriptive terms given by individuals, and to weight them so that the use of all seven points on a seven-point scale is at least a little less arbitrary.

One can, therefore, transfer the results of a technique such as method E to other subjective assessments, where more satisfactory verbal descriptions of the different qualities may be impossible, although such a transfer does involve making certain assumptions.

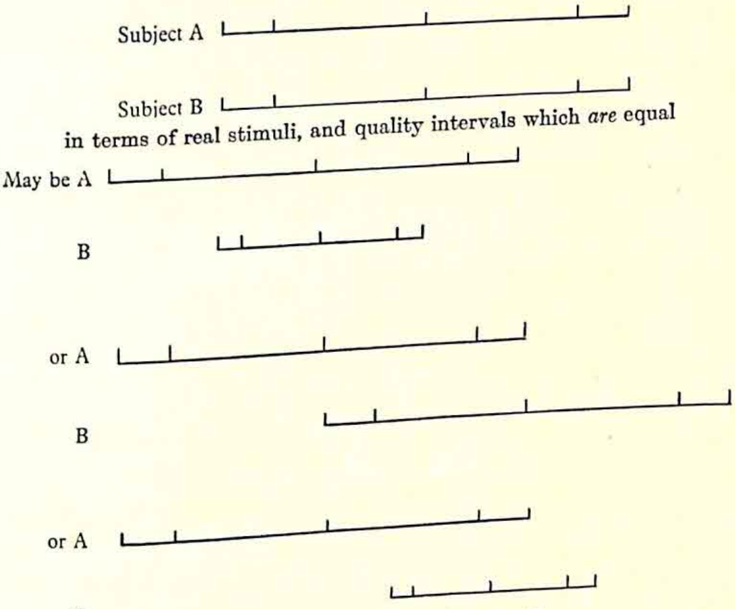
Thus, for example, it can only show the relative positions of the quality steps, one to another. There is no way of knowing from this whether subjects would agree about 'equal quality steps'; whether all judges would agree that a certain item is 'Very good', even if they agree to place 'Very good' in the same relative position to 'Extremely good'; or how long the scales are in relation to one another, even if the subjects agreed about 'equal quality steps'. Fig. 5 shows the sort of difficulty that is meant.

It might be possible to determine some of the answers to such problems by combining data obtained from a method such as test E with an examination of ratings on identical samples by different judges, though this would not be easy to do.

However, as previously stated, by choosing two values, such as 'Extremely good' and 'Extremely bad', and assuming that judges agree about these, one can at least make interpretation of subjective judgements one stage less arbitrary than by assuming that the qualities are equally spaced one from another. If nothing else, it reduces the number of assumptions and, apparently, one can improve the interpretation of subjective assessments by such methods.

It is clear, of course, that although such methods may make it possible to interpret the popular meaning of quantitative descriptive terms more successfully, it will still be necessary to select the other criteria, to which these terms will be applied, as carefully as possible (Hopkinson, 1950).

Fig. 5. Identical scales between arbitrary points, obtained by the method of test E


Subject A
Subject B
in terms of real stimuli, and quality intervals which *are* equal
May be A
B
or A
B
or A
B

Interpreting scales given for test E.

The results emphasize, however, that assessments made in terms of quantitative descriptions such as these should not be treated as if they were linearly related to one another. Furthermore, while it is true that subjects are to some extent suggestible, and will be affected by the order in which the descriptive qualities to be used in the rating scale are written out, since the only feature common to the results given for test E was that subjects inserted the phrases at positions which on the 'Bad' side, were a mirror image of those on the 'Good' side, it would seem better to construct rating scales so that the descriptive terms on each side of 'Average' also mirror each other.

SUMMARY

1. Some simple experiments were carried out to see how far people agreed in their use of adjectives qualified by quantitative adverbs, both for measurable items of everyday experience, and by a more abstract method.

2. The results suggest that if complete agreement, with no overlapping, is required, only three points on a 7-point rating scale of quality descriptions should be used.
3. A method is suggested by which the procedure of giving numerical values to these quality descriptions can be made one stage less arbitrary, and some experimental verification for such a method is quoted.

The author wishes to express his gratitude to his many subjects for their co-operation and, in particular, to the headmasters of the various schools for giving him permission to question their staffs. In addition, he wishes to acknowledge his debt to Dr G. W. Scott-Blair, Prof. R. C. Oldfield, Mr J. H. Prentice and Miss S. D. Carrinci for helpful discussions; to the latter for invaluable help in the analysis of the data; and to the Agricultural Research Council for a grant.

REFERENCES

HOLLINGWORTH, H. L. (1910). The central tendency of judgment. *J. Phil. Psychol. Sci. Meth.* VII, 461.

HOPKINSON, R. G. (1950). The multiple criterion technique of subject appraisal. *Quart. J. Exp. Psychol.* II, 124.

PRENTICE, J. H. (1953). The effect of storing and blending on the firmness and spreadability of butter. *Proc. XIII Dairy Congr.* II, 723.

SHEPPARD, D. (1953). Subjective assessments of firmness. The use of a rating scale. *Quart. J. Exp. Psychol.* V, 1.

TRESSELT, M. E. (1948). The effect of the experiences of contrasted groups upon the formation of a new scale of judgment. *J. Soc. Psychol.* XXVII, 209.

WOLTERS, A. W. P. (1916). The process of negation. *Brit. J. Psychol.* VIII, 183.

(*Manuscript received* 20 *August* 1953)

VERBAL EXTRAPOLATION AND INTERPOLATION

By E. C. POULTON

Medical Research Council Applied Psychology Research Unit, Cambridge

I. *Introduction* (p. 51). II. *Method* (pp. 51–52). III. *Scoring* (pp. 52–53). IV. *Results* (pp. 53–55). V. *Discussion* (pp. 56–57). VI. *Summary* (p. 57). *References* (p. 57).

I. Introduction

Experimental and statistical investigations of verbal redundancy (e.g. Miller, 1951; Newman & Gerstman, 1952) have used the method of extrapolation or forward-association. They have answered the question: 'How much new information is *added* by the next word or letter in a message?' This has been deduced from the frequency with which the component follows the previous one, two, three, ..., components. However, the method of interpolation or forward- and backward-association can also be used, and may in certain circumstances be more relevant, as it answers the question: 'How much new information is *given* by a particular component of a message?' This depends upon the frequency with which the component is found in its position in the particular type of passage, the component having to be related both to what has preceded it and to what follows it.

There are situations in everyday life permitting verbal extrapolation but not interpolation, as when a telephone conversation is cut off in the middle. But there are other situations permitting either, as when a sudden loud noise blots out the middle part of a message. Which of the two occurs in these circumstances can only be decided by experiment.

The original aim of the present experiments was to determine whether short gaps in verbal material are filled simply by extrapolation, or by interpolation. A subsequent aim was to discover whether by a suitable choice of the words immediately succeeding a short gap, interpolation could be made to give a wider range of answers than extrapolation.

II. Method

Extrapolation experiment. Of the thirty subjects, three were tested at a time, seated at tables screened from each other. A light short story of about 2000 words was played to them from a tape recorder. At 1 min. intervals the story was stopped. The thirty-one stopping places were selected on a time basis, independently of the content of the story. They were each marked by a sharp tap, and were followed by a 15 sec. pause. During the pause subjects had to write down the next two or three words of the story. They were instructed to answer as quickly as possible. Any answer was as good as any other, provided it fitted the context. Delays of more than 5 sec. in starting to write were noted by the experimenter, using a stop-watch. There was a short practice before the experiment, to ensure that subjects had understood the instructions.

Of the thirty-one stopping places, ten were at the start of a sentence, six of these after a link with the previous sentence such as 'and', the other four after the subject of the sentence; thirteen were after the subject and verb; and eight were after the subject, verb and object. The words immediately before a stopping place will be called the *precontext*.

Story-postcontext experiment. The procedure was similar for a new but comparable group of thirty subjects, except that they had to interpolate. The same thirty-one stopping places were each marked by a loud buzz lasting 1–2 sec. This was followed by a few more words taken from the story, the *postcontext*, before the 15 sec. pause. During the pause subjects had to write down the words omitted during the buzz.

The size of the actual gap left in the story ranged from one to eight words, the average being 4·1; but by reducing redundancy it could be filled fittingly by an average of 1·8 words. The start of the postcontext was chosen so that its meaning and grammatical structure affected the filling of the gap as little as possible. Generally it was a new sentence or a clause. In only two cases did the postcontext start after the subject of the sentence, so that this necessarily became part of the gap to be filled. The postcontexts ranged from 3 to 17 words, the average being 8·5.

Changed-postcontext experiment. In a second interpolation experiment nineteen of the postcontexts were altered in order to change the nature of the answers given by the group of thirty new but comparable subjects. A representative sample of twelve postcontexts was left unchanged. The mean postcontext length was reduced slightly to 7·9 words, the range being 2 to 16. The gaps could be filled by an average of 1·9 words.

Subjects. The ninety subjects were all naval ratings from Fleet Air Arm stations, aged between 17 and 31 years. Most had left school at 14 or 15 and had become apprentices or taken miscellaneous jobs before being called up for military service. Each performed the experiment once only.

III. SCORING

Commonest meaning changes. The number of extrapolation subjects giving the commonest meaning of their experiment at each stopping place, was compared with the number of subjects in the two interpolation experiments giving this meaning. Similarly, the number of story-postcontext subjects giving their commonest meaning (which might or might not be the same as for the extrapolation experiment) at each stop, was compared with the number from the changed-postcontext experiment. Classifying answers by meaning sometimes seemed quite arbitrary, especially at stops giving answers with many different meanings. In these cases meaning categories were widened until the commonest meaning of the first experiment was given by at least seven out of the thirty subjects (25 %), in order to ensure that the classification was reasonably representative of the answers given.

Changes in the incidence of commonest meanings in the comparison groups were predicted for one of three reasons: (*a*) the commonest meaning and the new postcontext meaning were in conflict (excessive redundancy, or contradiction); (*b*) the new postcontext demanded another meaning, the commonest meaning being unnecessary but possible in addition; and (*c*) the new postcontext demanded this meaning. In the latter case an increase was predicted, in the two former a decrease. These prediction criteria were formulated while comparing the extrapolation and story-postcontext answers. Only the other two comparisons are therefore shown in the results section below.

Changes in first word of answer. The same comparisons were made for commonest first words of answers as for commonest meanings. Changes in incidence were predicted using three similar criteria. Only the verbal structure of the postcontext was considered in this case, not its meaning. Where two first words were equally common in the first experiment, half weight was given to each. Nearly identical words, e.g. singulars and plurals, *a* and *an*, *round* and *around*, were grouped together. In the extrapolation-interpolation comparisons one stopping place had to be scrapped, as the precontext of the extrapolation experiment ended in the middle of a word, and was therefore changed for the two interpolation experiments.

Four-fifths of the commonest first-word predictions occurred at the same stops, and were in the same directions, as the commonest meaning predictions. The example which follows (stopping place no. 30) was chosen to show as many of the points mentioned above as possible:

Extrapolation experiment: 'but he wanted more (information), and he wheeled......'. Observed commonest meaning: on some member of the crew. Observed commonest first word: around (or round).

Story-postcontext experiment: 'and he wheeled......for the required illumination'. Observed commonest meaning: around in general, not on a member of the crew. Observed commonest first word: around (or round).

Changed-postcontext experiment: 'and he wheeled......of the cook for the required illumination'. Predicted changes in extrapolation experiment commonest meaning: decrease, because 'cook' in postcontext makes a member of the crew redundant. Predicted changes in extrapolation experiment commonest first word: decrease, because 'around' is unnecessary, while the new postcontext requires something like 'in the direction'.

Predicted changes in story-postcontext experiment commonest meaning: nil. Predicted changes in story-postcontext experiment commonest first word: decrease.

The change in the total number of different first words used at each stopping place was assumed to vary inversely with the change in the incidence of the commonest first word, and predictions were made accordingly. The contexts and predictions at the different stops are given in detail in Poulton (1953), together with a full analysis of the answers at one stop.

Indices of difficulty in answering. Omissions include failures to answer, and cases where subjects did not start to write until the story had recommenced after the 15 sec. pause. *Delays* are restricted to cases where subjects started to write between 5 and 15 sec. after the beginning of the pause. The number of answers which conflicted either in grammar or meaning with the precontext, or the postcontext when present, were also noted.

Calculations. Whenever possible the significance of the difference of each pair of means was assessed separately against the variability of the subjects for all stopping places in the group concerned, as well as against the variability of the stopping place differences for all the subjects combined. However, the significance levels of the latter comparisons have always been taken, as they were the less significant. Levels beyond $P=0{\cdot}001$ will be called 'highly significant', beyond $P=0{\cdot}01$ 'significant', and beyond $P=0{\cdot}05$ 'just significant'. Both tails of the assumed normal distribution have always been used.

IV. RESULTS

Extrapolation and changed-postcontext experiments

(1) The incidence of the commonest meanings of the extrapolation experiment is shown in Table 1 for this experiment and for the changed-postcontext experiment. Each stop has been allocated to one of three groups, depending upon whether its changed post-

context should have altered the incidence or not, and upon the predicted direction of the alteration, using the standard criteria given above. The differences between the two experiments and their p values are also shown.

(2) The incidence of the commonest first words of the extrapolation experiment is shown in Table 2 for the two experiments. Again each stop has been allocated to one of three prediction groups, but as compared with Table 1, a few of the stops have changed groups.

Table 1. *Extrapolation and changed-postcontext experiments: commonest extrapolation meanings*

	Changed postcontext predicted to produce		
	No change (17 stops)	Decrease (12 stops)	Increase (2 stops)
Extrapolation (%)	51	52	40
Changed-postcontext (%)	44	16	75
Difference (%)	7*	36	−35
P	0·05	0·001	0·05

* No change—decrease, $p = 0·001$; no change—increase, $p = 0·001$.

Table 2. *Extrapolation and changed-postcontext experiments: commonest extrapolation first words*

	Changed postcontext predicted to produce		
	No change (18 stops)	Decrease (11 stops)	Increase (1 stop)
Extrapolation (%)	45	46	33
Changed-postcontext (%)	39	18	50
Difference (%)	6	28*	−17
P	0·05	0·01	—

* No change—decrease, $p = 0·01$.

Table 3. *Extrapolation and changed-postcontext experiments: different first words*

	Changed postcontext predicted to produce		
	No change (18 stops)	Increase (11 stops)	Decrease (1 stop)
Extrapolation (%)	36	31	53
Changed-postcontext (%)	37	45	30
Difference (%)	−1	−14*	23
P	>0·05	0·05	—

* No change—increase, $p = 0·01$.

(3) The total number of different first words used in the extrapolation and changed-postcontext experiments is shown in Table 3. The prediction groups are identical with those of Table 2, but the increase and decrease predictions have been reversed.

(4) The changed-postcontext experiment gave significantly more delays in answering than the extrapolation experiment (14 % as compared with 5 %). It yielded significantly more answers which did not fit the precontext in grammar and/or meaning (4 % as compared with 1 %). In both cases the stops at which no change in answer was predicted showed just significant differences.

Story-postcontext and changed-postcontext experiments

(5) The incidence of the commonest meanings and commonest first words of the story-postcontext experiment is shown in Tables 4 and 5 for this experiment and for the

changed-postcontext experiment. The stops with postcontexts common to both experiments are shown separately, and the remaining stops have been allocated to one of two groups, depending upon whether the changed postcontexts should have decreased the incidence or not.

Table 4. *Story-postcontext and changed-postcontext experiments: commonest story-postcontext meanings*

	Same postcontext (12 stops)	Changed postcontext predicted to produce	
		No change (9 stops)	Decrease (10 stops)
Story-postcontext (%)	56	46	58
Changed-postcontext (%)	45	63	17
Difference (%)	11*	−17	41
P	0·01	0·05	0·001

* Same postcontext—no change predicted, $p=0{\cdot}001$; same postcontext—decrease predicted, $p=0{\cdot}001$.

Table 5. *Story-postcontext and changed-postcontext experiments: commonest story-postcontext first words*

	Same postcontext (12 stops)	Changed postcontext predicted to produce	
		No change (12 stops)	Decrease (7 stops)
Story-postcontext (%)	44	42	55
Changed-postcontext (%)	37	38	26
Difference (%)	7	4	29*
P	>0·05	>0·05	0·05

* Same postcontext—decrease predicted, $p=0{\cdot}05$.

Table 6. *Story-postcontext and changed-postcontext experiments: different first words*

	Same postcontext (12 stops)	Changed postcontext predicted to produce	
		No change (12 stops)	Increase (7 stops)
Story-postcontext (%)	37	37	27
Changed-postcontext (%)	38	38	44
Difference (%)	−1	−1	−17*
P	>0·05	>0·05	0·01

* Same postcontext—increase predicted, $p=0{\cdot}05$.

(6) The total number of different first words used in the story-postcontext and changed-postcontext experiments is shown in Table 6. The prediction groups are identical with those of Table 5, the predicted decrease having become a predicted increase.

(7) The changed-postcontext experiment gave significantly more delays than the story-postcontext experiment (14 % as compared with 7 %). It also yielded just significantly more answers not fitting the postcontext (17 % as compared with 9 %). Answers not fitting the precontext remained unchanged.

(8) Omissions showed no significant differences at all, always lying in the region of from 2 to 4 %.

V. Discussion

The influence of postcontext. In Tables 1 and 2 the extrapolation and changed-postcontext experiments have been compared on the commonest meanings and commonest first words of the extrapolation experiment. This experiment would be expected to yield a slightly higher incidence than the other by chance alone (just as a bridge player's longest suit in one hand tends to be longer than this same suit in his next hand). This expectation is borne out by the scores at the stops where no change was predicted, for in both tables the extrapolation experiment shows a just significant advantage. It distorts direct comparisons between the two experiments at the stops at which changes were predicted. However, the differences between the experiments at the predicted no-change and predicted decrease and increase stops can be compared, because this trend should affect all three groups of stops about equally. In both Tables 1 and 2 the relationships between these differences are as predicted, the p values being given as footnotes.

A direct comparison can be made between the extrapolation and changed-postcontext experiments in the total number of different first words used. Table 3 shows that there was no difference at the stops at which no change was predicted, but differences as predicted at the remaining stops.

Tables 1–3 thus show that changing a sentence continuation task from extrapolation to interpolation can alter the answers given, provided the postcontexts are suitably chosen as they were at the prediction stops of the changed-postcontext experiment. Result 4 shows that even when the postcontext did not alter the answers appreciably, it tended to add to the difficulty of answering, thus indicating that some attention was being paid to it. Clearly, when a short section of a verbal message is lost in transmission, the gap will be filled by reference to what follows it as well as to what precedes it. The forward-association method of study (Miller, 1951, pp. 187–8; Newman & Gerstman, 1952) would thus be an over-simplification of this situation.

The influence of the postcontext is also shown in Tables 4–6, in which the original story postcontexts are compared with the changed postcontexts. In tables 4 and 5 it is the differences between the story-postcontext and changed-postcontext experiments which should be compared (cf. discussion of Tables 1 and 2 above). In both tables the relationship between the same postcontext and predicted decrease groups of stops is as predicted (see p values in table footnotes). In Table 6 a direct comparison can be made between the two experiments, and again the difference is as predicted.

The high incidence of the commonest story-postcontext meanings found in the changed-postcontext experiment at the stops at which no change was predicted (Table 4) requires comment. At seven of the stops the changed postcontext appeared to suggest this meaning more strongly than the corresponding original story postcontext. However, none of the differences in postcontext was sufficiently definite for prediction to be possible using the standard criteria given above.

Context and range of answer. Many of the changed postcontexts were designed not to fit the obvious meanings and/or first words suggested by the precontexts. At these stops the changed-postcontext experiment yielded less uniformity in answer than the extrapolation experiment. Table 3 shows the resultant increase in the number of different first words used (predicted increase stops). The decrease in the commonest extrapolation meanings (Table 1 predicted decrease stops) suggests that there was probably a corresponding increase in the range of meanings, although no satisfactory measure of this could be devised. This reduced uniformity was not due simply to an increase in answers not fitting the precontext, for the latter was small and was spread more or less evenly over all the stops (Result 4).

In these cases the additional constraint imposed by the postcontext increased the range of the answers. Thus by *adding* to a context it is possible to *reduce* its effectiveness in sign-posting the answers. This emphasis on reduced uniformity of answer, rather than on increased uniformity in the design of the changed postcontexts, must account for the greater difficulty of this experiment as compared with the story-postcontext experiment (Result 7).

VI. SUMMARY

A comparison was made between continuing sentences (extrapolation) and filling gaps of unstated size in sentences (interpolation), when both operations occurred at the same randomly selected places in a story. It was found that the operations did not necessarily give different answers, although interpolation was more difficult than extrapolation. However, by a suitable choice of postcontexts, changes could be produced in the meanings and first words of the answers. The addition of a postcontext could actually increase the range of the answers given.

The influence of Prof. Sir Frederic Bartlett, F.R.S., is clearly shown in this paper. Statistical advice was received from Mr N. L. Webb. The subjects were supplied by the Royal Navy. Financial support from the Medical Research Council is also gratefully acknowledged.

REFERENCES

MILLER, G. A. (1951). *Language and Communication*. New York: McGraw-Hill.
NEWMAN, E. B. & GERSTMAN, L. J. (1952). A new method for analysing printed English. *J. Exp. Psychol.* XLIV, 114–25.
POULTON, E. C. (1953). Verbal extrapolation and interpolation. *Cambridge Appl. Psychol. Unit*, Rep. no. 172.

(*Manuscript received* 5 *September* 1953)

A PRELIMINARY INVESTIGATION OF THE EFFECTS OF CONTINUOUS STIMULATION ON THE PERCEPTION OF 'APPARENT MOTION'

BY M. B. SHAPIRO

Department of Psychology, Institute of Psychiatry, Maudsley Hospital

I. *Introduction* (p. 58). II. *Explanatory hypotheses and findings* (pp. 58–60). III. *Deductions from the hypotheses* (p. 60). IV. *The experiment* (pp. 60–62): (i) *Purpose* (pp. 60–61); (ii) *Apparatus* (p. 61); (iii) *Sample* (p. 61); (iv) *Procedure* (pp. 61–62). V. *Results* (pp. 62–64). VI. *Discussion* (p. 64). VII. *Alternative explanations of the findings* (pp. 64–67): (i) *Displacement effect* (pp. 64–65); (ii) *A lowered intensity of apparent stimulation* (pp. 65–67). VIII. *Summary* (p. 67). *References* (p. 67).

I. INTRODUCTION

The experiment reported in this paper is part of an investigation into the psychological effects of brain damage. The particular observation which was the starting-point for this experiment was one reported by Werner & Thuma (1942) who found that, under certain conditions, brain-damaged subjects failed to see apparent motion.[1] The conditions consisted of presenting the subjects, by means of a modified Dodge tachistoscope, with successive exposures of two vertical black lines. At a certain speed of succession all but one of twenty normal subjects saw movement. Only two of twenty brain-damaged subjects saw movement.

Both groups were of low I.Q., the average of the brain-damaged group being 69, and of the non-brain-damaged group being 72. It is stated that the brain-damaged group did not suffer from any gross motor defects or visual agnosias, although specific diagnoses of the type of brain damage suffered by these subjects are not stated.

A related finding is that of Bender & Teuber (1949), who found that patients with occipital-parietal lesions failed to perceive apparent motion when certain affected parts of the visual field were appropriately stimulated.

II. EXPLANATORY HYPOTHESES AND FINDINGS

Two hypotheses were formulated in explanation of this finding:

(i) The first was in essence the same as Wertheimer's explanation of the occurrence of apparent motion in normals. According to this hypothesis visual stimulation in any relevant part of the brain is followed by the irradiation of excitatory effects from the site of stimulation. If such stimulation is followed by stimulation in another part of the brain, sufficiently soon after and appropriately placed, the irradiation effect of the second site of stimulation is presumed to combine with the irradiation effect of the first in some way so as to produce the visual perception.

It should be noted that it is not the full and detailed Wertheimer hypothesis which is being formulated here. All that is being used is the general notion of the irradiation of excitatory effects and their combination to provide the basis of the perception of apparent motion.

---

[1] This finding was brought to my attention by Mrs M. Brenner.

Some experiments have produced results which appear to throw considerable doubt on this theory. Closer examination of these experiments shows, however, that they either are based on incorrect assumptions or deal only with secondary aspects of the Wertheimer hypothesis.

Two examples will be given in illustration. The first of these is an article by K. U. Smith (1940). He found that the same nystagmic head movements were produced in the guinea-pig by real motion and 'apparent motion'. These movements were not affected by the removal of the animal's cortices. This experiment, Smith believes, demonstrates the inadequacy of the Wertheimer theory. He was apparently unaware of the fact, discussed by Marquis (1935), that in the lower vertebrates the visual functions are carried out entirely by the mesencephalic centres, while in man the cortex has taken over all the optic functions with the exception of the pupillary reflex to light.

Our second example of a claimed disproof of the Wertheimer hypothesis is a paper by K. R. Smith (1948). Ten human subjects were given successive visual stimulation so that each stimulus fell upon the retina of its corresponding eye at the nasal end of the horizontal meridian. The visual projection areas of these two parts are completely separated from each other by the longitudinal fissure of the cerebrum, one lying on the right and the other on the left occipital lobe. The afferent fibres leading to these projection areas are also separate. Thus the two experimental stimuli were represented respectively at the extreme margin of the left visual cortex and at an opposing point on the right visual cortex. Seven of the ten subjects reported complete horizontal movement. Smith concludes from this that the general theory of corresponding interaction is thrown into considerable doubt. The first point to be noticed is that none of Smith's subjects reported the unequivocal perception of object movement; only 'pure phi' was definitely seen. 'Pure phi' is an experience of movement in which no moving object is seen. The second objection to Smith's conclusion is that a connexion between the two occipital poles may exist via the corpus callosum and possibly the lateral geniculate bodies. Any two parts of the brain which are interconnected may interact.

In view of such considerations the necessity still remains for testing the Wertheimer hypothesis in the generalized form described above.

(ii) The second hypothesis was that one of the effects of brain damage on psychological processes is the increase of inhibitory effects in the affected area. Inhibitory effects are here defined as the demonstrable decrease or disappearance of excitatory effects as the result of some process in the brain. The intensity and range of inhibition will vary, in as yet undefined ways, with the site and extent of the lesion. The grounds for this hypothesis were the work of Teuber & Bender (1949) and Pavlov (1927).

The findings of Teuber & Bender (1949) were obtained on patients who had occipito-parietal lesions. They report that extinction occurs in mildly impaired areas of the visual field, as soon as the other relatively intact areas are stimulated simultaneously. Thus a pattern may be seen on exposure to an amblyopic half of a field, provided this exposure takes place against a relatively homogeneous background. However, as soon as the second pattern is simultaneously exposed in the other (less impaired) half of the field, the first pattern becomes 'extinct'. 'The patient states, often with signs of surprise, that he does not see the first pattern any longer. . . . However, as soon as the pattern in the less impaired part of the field is removed, the one in the affected field becomes visible.'

Pavlov's findings were obtained from dogs which had had part of their cortices removed.

For a period after operation, previously established conditioned reflexes related to the site of the injury could not be evoked. Instead, a definite inhibitory effect in the form of sleep was produced.[1]

What is common to both these findings is that, in descriptive terms, inhibitory effects are increased as a result of brain damage. We are not concerned here with the mechanisms underlying these findings about which there is some disagreement but only with the facts at a descriptive level.

III. DEDUCTIONS FROM THE HYPOTHESES

From these two hypotheses it would follow that one possible explanation of the failure of the brain-damaged patient to see apparent motion is that the irradiation effects postulated by the first hypothesis above are inhibited, thus removing the basis for the perception of apparent motion.

One deduction which follows from this explanation is that the introduction of an inhibitory condition in normal subjects should result in their failing to see apparent motion at speeds of succession which normally would be sufficient. No assumption is being made here about the nature of the inhibitory processes involved. We are dealing here only with effects. Such a condition should result at least in the need to shorten the time interval between the two stimulations in order to produce the perception of motion.

In order to test this deduction it was necessary to find a method of producing an inhibitory condition in the brains of normal subjects. It was decided to do this by the method of continuous visual stimulation. This was suggested by the work of Kohler & Wallach (1944), who report that continuous visual stimulation can result in making the stimulus phenomenologically less intense. Such stimulation also has inhibitory effects on the surrounding areas of the brain, the effect decreasing as the distance from the point of stimulation increases. These facts have been demonstrated by Bevan (1951). He investigated the changes in the intensity thresholds of the perception of a light stimulus at varying distances from the area of continuous stimulation. He used a vertical line of light as his continuous stimulus, and a dot of light as his test stimulus. He first of all established the intensity threshold for a dot of light at a given position. This was followed by stimulation with the vertical line for 20 sec. This in turn was followed by a re-determination of the threshold. This procedure was repeated with dots of light at different distances from the vertical line.

Two findings emerged. The first was that continuous stimulation resulted in a rise of thresholds. The second was that the rise in threshold decreased as the distance from the site of continuous stimulation increased.

IV. THE EXPERIMENT

(i) *Purpose*

To test the deduction that normal subjects would be unable, after they had been subjected to a period of continuous visual stimulation, to perceive visual apparent motion at speeds of successive stimulation which were normally fast enough to evoke such perception.

---

[1] A re-reading of Pavlov since the completion of this experiment shows that the long-term effects of brain damage on inhibitory effects were rather to weaken than to intensify them. This poses further problems which are not relevant to the outcome of this particular paper, such as the possibility of differences between lower and higher animals.

(Mrs M. Brenner was responsible for the final details of the experiment and for carrying out the testing.)

(ii) *Apparatus*[1]

The apparatus consisted of a modified Dodge tachistoscope copied from the model given by Werner & Thuma (1942) in their article; the timing device was electronic, and had continuously variable speeds for the time interval from 0·110 to 0·86 sec. Exposure time was kept constant at 0·09 sec.

The apparent motion figures consisted of two white lines, $\frac{1}{20}$ in. wide and $\frac{5}{8}$ in. long, making an angle of 70°. The lines were made of white sellotape mounted on black velvet cards, 6 × 6 in., and viewed at 20 in. from the subject. The lines were so placed that the one side of the angle was horizontal, and the angle pointed to the left (see Fig. 1).

For continuous stimulation it was decided to use a figure consisting of a black circle, 2 in. diameter, mounted on a white screen 3 × 3 ft. It stood at eye level, 6 ft. away from the subject, so as to occupy the same visual angle as the two lines. The area of continuous stimulation was thus placed in the path of the apparent motion.

A chin rest was used during fixation, and the experiment was carried out in a semi-darkened room. To avoid complication from the after image, and to make sure we were dealing with central and not peripheral functions, the continuous stimulation and apparent motion were done monocularly, using a different eye for each—the dominant eye being used for tests of apparent motion.

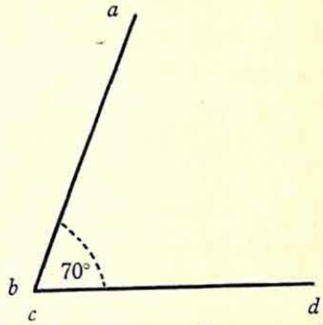


Fig. 1. Line *ab* was first exposed, and followed by line *cd*.

(iii) *Sample*

The subjects for this experiment were thirty-three normal intelligent adults. Of the thirty-three subjects, seven failed to report movement—so the complete experiment used twenty-six subjects who were divided into an experimental and a control group. Table 1 gives the distribution of the groups with respect to age, sex, visual acuity and sophistication in experiments of a visual perceptual nature.

Table 1. *Distribution of age, sex, vision and sophistication of two groups of subjects*

	Age		Sex		Vision			Sophistication	
	21–25	25–35	M.	F.	Normal	Corrected	Uncorrected	No experience	Some experience
Experimental	6	7	6	7	10	2	1	8	5
Control	5	8	7	6	8	5	0	9	4

(iv) *Procedure*

Test 1. First the threshold for apparent motion for each subject was established on five trials using the limiting method. The lower threshold was the average of the longest time intervals at which movement was reported for each of the five trials.

In preliminary trials many subjects reported difficulty in deciding whether or not they saw clear optimal movement. Subjects were therefore instructed to answer 'no' when they were quite sure that they perceived clear succession ('two separate lines one after the other') and to answer 'yes' to 'any impression that the first line moved down to the place of the second'.

---

[1] Mr W. Withers constructed the timing device and Mr J. Westhenry the tachistoscope.

Test 2. A second test was carried out on each subject between 24 and 48 hr. later. In this test, the experimental group was first treated with continuous stimulation in the appropriate part of the visual field. This was done by fixating a cross in the centre of the circle described above for 3 min. The subjects kept as still as possible while doing this. Immediately afterwards, the threshold for apparent motion was established in the same way as in Test 1, with the same instructions. The time between the two procedures was not measured, but the one followed immediately upon the other in the literal sense of the word.

The subjects of the control group had their thresholds for apparent motion established without prior fixation.

It is to be noted that, although several subjects commented on the difficulty of deciding whether or not they saw movement, the thresholds appeared to be relatively consistent for the five trials of each test. This might be due to the instructions which resulted in any experience which was not definitely 'successive' being regarded as 'apparent movement'.

V. RESULTS

Table 2 gives the lower thresholds of each subject for each of the five trials in the first and the second tests. As it was not possible to produce simultaneity under the given conditions, changes in the upper thresholds could not be measured and are therefore not given.

The results are consistent with expectation. The lower thresholds for the experimental group rise, i.e. after continuous stimulation the subjects fail to see apparent motion at speeds of succession which were sufficient in the first test. They now require higher speeds. The difference between the first test and the second test gives a correlated t of 3·57 which is significant at the 0·005 level. The mean lower threshold of the control group also rises, but insignificantly, giving a correlated t of 0·64. (See Table 4.)

To check the possibility of abnormal sampling influencing the result, the mean difference between the experimental and control groups at the first test was examined. The control has a mean threshold which is higher than that of the experimental group, but not significantly so.

The difference between the control group and the experimental group on the second test is even smaller and the t falls to 0·84 (Table 3). To analyse this finding further the mean threshold for each of the five trials run in each test was calculated and the results plotted in the graph in Fig. 5, p. 66.[1] Two main findings emerge.

First, in the experimental group there is a marked rise in threshold in the trial immediately after the continuous stimulation, and this threshold falls steadily except for the fourth trial which is roughly the same as the third. That this reflects a real trend is shown by the fact that the first two trials of the first test differ at the 0·005 level of significance from the first two trials of the second test. However, the last trials of each test do not differ significantly. This applies only to the experimental group. There is no such trend in the control group (Table 4).

The second finding is that the mean threshold for trial a of the second test is higher in the experimental group than in the control. The difference is significant at the 0·02 level.

---

[1] This was done at the suggestion of Mrs M. Brenner.

On trial *b*, which is the second trial, the difference between the two groups is not significant (Table 3).

It would seem, therefore, that the results are in every way consistent with expectation. This is confirmed by comparing the mean change from test 1 to test 2 in the experimental group with the same change in the control group. The change in the experimental group is larger and the difference significant at the 0·005 level (see Table 3).

Table 2. *Showing all the lower thresholds obtained in milliseconds. Blanks are shown where the subject said he saw nothing*

Control group

	Test 1. Trials						Test 2. Trials					
S	*a*	*b*	*c*	*d*	*e*	Mean	*a*	*b*	*c*	*d*	*e*	Mean
1	120	400	300	300	300	284	300	120	400	120	300	248
2	400	120	200	300	300	264	300	200	200	300	300	260
3	300	150	150	150	200	190	200	200	150	300	150	220
4	110	300	120	150	120	160	200	200	200	150	150	180
5	110	200	150	200	200	172	120	200	200	200	200	184
6	200	120	150	200	150	164	200	150	150	200	150	170
7	—	150	—	150	110	136·67	200	200	150	150	150	170
8	200	150	300	200	300	230	200	200	200	300	200	220
9	200	200	150	200	200	190	120	150	120	200	200	158
10	300	200	200	300	300	260	200	200	300	200	200	220
11	200	200	120	—	—	173·33	120	120	120	120	120	120
12	500	200	400	500	400	400	400	500	300	500	500	440
13	300	200	400	200	300	280	200	300	300	200	200	240
Mean	245	199·23	220	237·5	240		212·31	210·77	214·62	233·85	216·92	

Experimental group

	Test 1. Before continuous stimulation						Test 2. After continuous stimulation					
S	*a*	*b*	*c*	*d*	*e*	Mean	*a*	*b*	*c*	*d*	*e*	Mean
1	300	200	300	200	200	240	150	200	200	300	200	210
2	300	120	200	110	200	186	120	150	120	150	120	132
3	150	150	120	150	200	154	110	110	120	200	400	188
4	200	200	300	300	300	260	150	150	200	150	400	210
5	400	400	200	500	300	360	150	120	500	200	400	274
6	200	200	200	150	200	190	150	300	200	300	300	250
7	300	300	400	300	400	340	200	200	300	150	300	230
8	300	200	300	300	200	260	150	150	—	200	200	175
9	400	400	300	300	300	340	150	400	200	200	300	250
10	150	150	300	200	150	190	110	150	200	—	200	165
11	150	400	300	300	200	270	—	200	200	200	200	200
12	200	200	300	200	200	220	—	120	150	150	150	142·5
13	200	500	300	200	200	280	120	120	120	150	120	126
Mean	250	263·08	270·77	246·92	234·62		141·82	182·31	209·17	195·83	253·08	

Table 3. *Comparison of control group with experimental group*

Measures compared	Control		Experimental		Difference	t* (Uncorrelated)	Level of significance P	D.F.
	Mean	S.D.	Mean	S.D.				
Means of test 1	223·38	69·19	253·08	62·74	29·70	1·10†	N.S.	24
Means of test 2	217·69	74·86	196·39	45·38	21·30	0·84	N.S.	24
Trial '*a*' on test 2	212·31	77·07	141·82	24·79	70·49	2·78	<0·01	22
Trial '*b*' on test 2	210·76	94·60	182·31	80·01	28·45	0·80	N.S.	24
Change from test 1 to test 2	5·69	30·63	56·69	54·97	51·00	2·81	<0·005	24

* Formula for Uncorrelated 't': $t = \frac{M_1 - M_2}{\sqrt{\left[\frac{\Sigma x_1^2 + \Sigma x_2^2}{(N_1 + N_2) - 2}\left(\frac{1}{N_1} + \frac{1}{N_2}\right)\right]}}$.

† Two-tail test; all other t-tests one-tail.

We can conclude, therefore, that the data as they stand are consistent with the notion that continuous stimulation has an immediate effect on threshold for the perception of apparent motion but that effects last for a short time.

VI. DISCUSSION

These findings are consistent with those reported by Detherage & Bitterman (1952), whose paper appeared while the experiment reported in this paper was in progress. They do not give details of their results, but they report an increase in the threshold of apparent motion resulting from a period of continuous stimulation, and indicate the connexion between this result and those obtained by Teuber & Bender (1949) with brain-damaged patients.

Table 4.

Measures compared	Groups	Test 1 Mean	Test 1 S.D.	Test 2 Mean	Test 2 S.D.	Difference	t* (correlated)	Level of significance P	D.F.
Test 1 and test 2	Control	223·38	69·92	217·69	74·86	5·69	0·64	N.S.	24
	Exptl.	253·08	62·74	196·39	45·38	56·69	3·57	0·005	24
Means of trials $a+b$ of test 1 with means of trials $a+b$ of test 2	Control	218·46	53·84	211·54	77·81	6·92	0·51	N.S.	24
	Exptl.	256·54	81·65	163·46	47·05	93·08	4·26	0·005	24
Means of trials $d+e$ of test 1 with means of trials $d+e$ of test 2	Control	233·69	84·97	225·38	92·78	8·31	0·55	N.S.	24
	Exptl.	240·77	73·27	224·62	58·37	16·15	0·75	N.S.	24

* Formula for Correlated 't':

$$t=\frac{\bar{d}}{\sqrt{\left[\frac{\Sigma d^2-\frac{(\Sigma d)^2}{N}}{(N_1-1)\,N_1}\right]}}.$$

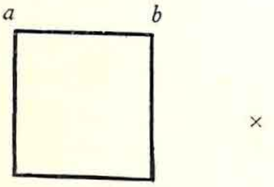


Fig. 2.

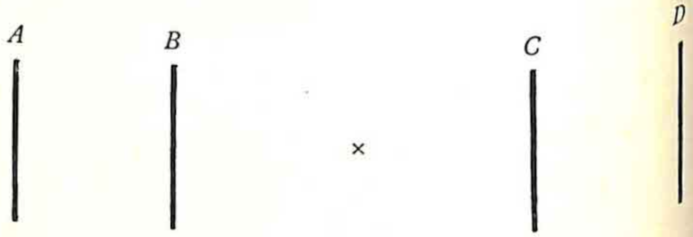


Fig. 3.

VII. ALTERNATIVE EXPLANATIONS OF THE FINDINGS

While the results reported above are consistent with the explanatory hypotheses advanced, further consideration shows that there are at least two alternative hypotheses which will account for the results as well as the irradiation theory.

(i) *Displacement effect*

One of the demonstrated effects of continuous stimulation is the displacement effect. As an example, let us assume that an area of the visual field has been under stimulation for 2 min., by a square figure (see Fig. 2). In the illustration × marks the point of fixation. At the end of the 2 min., let us assume, the square figure is removed and the two pairs of lines as shown in Fig. 3 are exposed. In this illustration × is again the fixation point. The distance between lines A and B is equal to the distance between lines C and D, the distance between line B and the fixation point is the same as the distance between line C and the fixation point. The distance between A and B is just greater than the distance between a and b of the square figure shown in Fig. 2. When Fig. 2 is exposed the distance between *A* and *B* will appear to be greater than the distance between *C* and *D*. This is the kind of displacement effect which results from continuous stimulation. This effect was the main feature of what Kohler & Wallach (1944) called the 'figural after-effect'. It was the central point of their investigations, and was demonstrated repeatedly in different ways.

A related effect has been demonstrated in brain-damaged patients by Teuber & Bender (1949), who report that when a moving light was brought in a straight line tangentially towards an area affected by brain damage determined by perimetry, it appeared to curve away from that area. That this may be due to a displacement effect, similar to that reported by Kohler & Wallach, is shown by the work of Detherage & Bitterman (1952) with normal subjects. They found that the stroboscopic stimulus appeared to curve away from the area which had been under continuous stimulation (see Fig. 4), just as the moving light appeared to curve away from the affected area in brain-damaged patients.

When we come to examine the present experiment, it is clear that the continuous stimulation in the second test might have increased the apparent distance between the two lines. This might have been responsible for the rise in threshold of apparent motion.

(ii) *A lowered intensity of apparent stimulation*

A second effect of continuous stimulation is the one described by Bevan (1951) and referred to above, in which the luminosity was lowered by a previous period of continuous stimulation. This alone might be sufficient to raise the thresholds for apparent motion.

The formulation of these two theories shows that while we have results which are consistent with a generalized Wertheimer hypothesis, the alternative explanations still remain to be tested.

It might be thought that we have in the displacement effect an alternative to our second hypothesis: that inhibitory effects are responsible for raising the threshold for the perception of visual apparent motion. Consideration of the relevant data in the literature shows, however, that the kind of displacement effects necessary to raise the threshold have only been observed in association with inhibitory effects.

There appear to be two kinds of displacement effect. First of all there is the displacement effect which was described in section III, in which the co-existence of an inhibitory and an excitatory effect results in the displacement of the latter *away* from the former. Secondly, there is the displacement effect to be observed when, under certain conditions, two excitatory effects coincide. Under these conditions the two excitatory effects appear to be displaced *towards* each other. For example, Kohler (1940) reports that when two parts of the body are touched simultaneously, the stimulations will, under certain conditions, appear to be displaced towards each other. He reports this phenomenon for touch, vision, and sound. Thus it seems that the kind of displacement effect which will raise the threshold for apparent motion can only appear in the presence of an inhibitory effect. We do not say, at this stage, by what mechanism this displacement takes place.

In evaluating these results it must be remembered that many of the subjects had difficulty in deciding whether or not they saw clear optimal movement, and in fact seven of the thirty-three subjects tested failed to see any movement. It was decided to continue the experiment because no reason could be seen why continuous stimulation should not have the predicted effect on any form of apparent motion. On the other hand, the

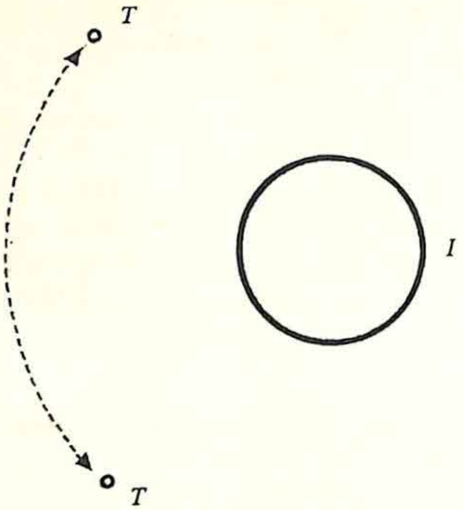


Fig. 4. Taken from Detherage & Bitterman (1952). *I* = inspection figure; *T* = test figure (alternating interocular presentation). All figures appear as light areas against a dark ground.

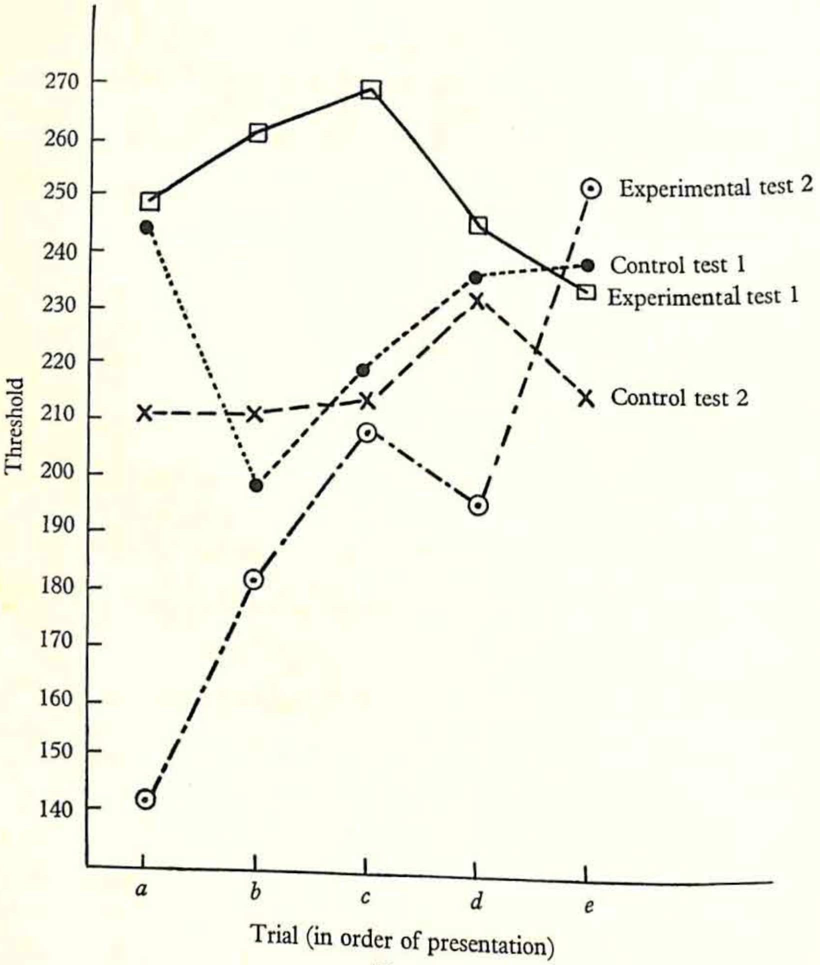


Fig. 5.

possibility still exists that where object movement is clearly seen, continuous stimulation would cease to have an effect.

VIII. SUMMARY

An explanation of the reported inability of some brain-damaged subjects to see apparent motion was advanced. This explanation was that there was an exaggeration of inhibitory effects in brain-damaged patients, and that these effects weakened or destroyed the irradiation of excitatory processes, which was postulated as the basis of the perception of apparent motion. This explanation was tested by inducing inhibitory effects by means of continuous stimulation in the appropriate area of the visual field of normal subjects. The results show a significant rise in the threshold in the experimental group, as was predicted.

Thanks are due to Mr A. Yates for the statistical analysis and criticisms, to Dr J. Brierley, Senior Lecturer of the Department of Neuropathology, for discussion and advice and to Miss A. Davis for typing and preparing the manuscript. Part of the work reported in this paper was made possible by a grant from the Research Fund made available from the endowment by the Board of Governors of the Bethlem Royal Hospital and the Maudsley Hospital.

REFERENCES

BENDER, M. & TEUBER, H. L. (1949). Disturbance in visual perception following cerebral lesions. *J. Psychol.* XXVIII, 223–33.

BEVAN, W., Jr. (1951). The influence of figural after-effects upon visual intensity thresholds. *J. Gen. Psychol.* XLV, 189–207.

DETHERAGE, B. H. & BITTERMAN, M. E. (1952). The effect of satiation on stroboscopic movement. *Amer. J. Psychol.* LXV, 108–9.

KOHLER, W. (1940). *Dynamics in Psychology*. New York: Liveright Publishing Co.

KOHLER, W. & WALLACH, W. (1944). Figural after effects. An investigation of visual process. *Proc. Amer. Phil. Soc.* LXXXVIII, 269–357.

MARQUIS, D. G. (1935). Phylogenetic interpretation of the functions of the visual cortex. *Arch. Neurol. Psychiat.* XXXIII, 807–15.

PAVLOV, I. (1927). *Conditioned Reflexes*. London: Oxford University Press.

SMITH, K. R. (1948). Visual apparent motion in the absence of neural interaction. *Amer. J. Psychol.* LXI, 73–9.

SMITH, K. U. (1940). The neural centres concerned in the mediation of apparent movement vision. *J. Exp. Psychol.* XXVI, 443–63.

TEUBER, H. L. & BENDER, M. (1949). Alterations in pattern vision following trauma in the occipital lobes in man. *J. Gen. Psychol.* XL, 38–57.

WERNER, H. & THUMA, A. (1942). A deficiency in the perception of apparent motion in children with brain injury. *Amer. J. Psychol.* LV, 58–67.

(*Manuscript received 22 July* 1953)

CORRESPONDENCE

THE EDITOR,
British Journal of Psychology.

Dear Sir,

As I have recently been supervising some research into the problems of chromatic aberration I would like to point out the relevance of this well-established phenomena as applied to Dr Zająk's paper[1] describing the results of an investigation of the influence of colour on depth perception. Some years ago I examined the effect of various colours on depth perception, but did not report the findings as they merely added to the general negative nature of the results concerning the differential effects of monochromatic light on visual acuity. However, the theory of chromatic aberration is applicable where objects in the visual field are variously illuminated with different colours. Thus, Zająk comments that 'a relation seems to exist between depth perception and wave-length of light' and 'objects viewed in indirect vision through different colour filters are seen at different distances from the eyes'.

In the second chapter of his recent text, Hartridge discusses the phenomena of chromatic stereoscopy, pointing out that if the eyes are fixated on a point the violet rays coming from that point will be most refracted and hence will appear to come from a more distant object than the red rays. Thus, if we were to aline objects, in the one case comparing a red with a white object, and, in the other, a blue with a white object, then the distribution of settings in the blue would fall closer to the observer than that for the red, i.e. the blue rod, appearing further due to chromatic differential refraction, will be set nearer to the observer for equidistance, and vice versa for the red. Thus, Zająk writes that 'objects are seen as nearer with longer wave-lengths and as further away with shorter wave-lengths of light' but nowhere indicates the well-known basis of this fact.

It must be further indicated that accurate measurements will not be possible with a lateral mounting such as he has used, due to the effects of the eccentric situation of the fovea in respect of the optical axis of the eye. An experiment which used vertically mounted rods, differentially illuminated, would be a useful addition to knowledge in this field as the results could then be directly compared with those of Berry *et al.*, using the stereo-vernier apparatus suggested by Walls.

The excellent theoretical analysis of depth perception by Walls in terms of summation of subliminal monocular cues is not referred to by Zająk, but in my opinion the greatest defect in this paper is the omitting of well-established physiological findings. Experimenters in this field should also come to some agreement as to the best procedures so that results could then be compared directly and not, as at present, by analogy.

Yours faithfully,

EDGAR HOWARTH
Lecturer in Psychology

Department of Psychology
The University of Melbourne
Australia
23 *June* 1953

---

[1] Zając, J. L. (1953). Investigations on depth perception in indirect vision. *Brit. J. Psychol.* XLIV, 132–44.

THE EDITOR,
British Journal of Psychology.

Dear Sir,

I read with interest Dr Howarth's letter to you about my paper. If I understand the meaning of his observations, his remarks refer mainly to the relation between the chromatic aberration and depth perception in *direct* and *binocular stereoscopic* vision, while the object of my investigations was the depth perception in *monocular indirect* and *binocular indirect vision with double images.* My hypothesis about differences in depth perception in monocular indirect vision resulting from differences in wave-length of light rays was advanced as early as 1913. In my recent paper I mentioned the Matthiessen formula which in my opinion is a very good illustration of the 'well-established physiological findings' as well as of other physical data mentioned by him.

I also referred to Hartridge's discussion of chromatic stereoscopy mentioned by Mr Howarth, pointing out, however, that it is connected with direct vision.

I also mentioned three different effects of refraction in the eye of pencils of light of different wave-lengths, i.e.: (1) the difference in size of the images on retinas, (2) the difference in distinctness, (3) the difference of lateral distance from the image of the fixation point; which all are results of difference of focus within the eye of an image of an object in lateral field of vision, and are related to changes of accommodation of the eye. All these effects are also undoubtedly 'physiological findings'.

I am grateful to Mr Howarth for mentioning one more theoretical analysis of depth perception by Walls, but if it refers to the stereoscopic vision, it has nothing directly to do with my investigations. Neither do the phenomena discussed by Hartridge as quoted by Mr Howarth in his letter.

I have now studied the influence of colour on depth perception in stereoscopic vision, and the preliminary results seem to confirm the fact, that, in stereoscopic vision with colours as well as without, the principal factors involved are different from those in monocular vision or in binocular vision with double images, although relations can be established between those two kinds of perception.

It must be strongly emphasized that a distinction should be made between what we *see* about depth and what we *estimate* about the distance of objects. In the case of direct monocular or binocular fixation of a single object in the field of vision there is more a question of *estimating* the distance at which that object is placed than of *seeing* in depth. In my opinion it does not make sense to say that we see an object at a distance, at which it is placed in 'reality' or geometrically. But we can make estimates (rightly or wrongly) about the distance at which an object is placed geometrically. The failure of many authors to distinguish between those two notions has led to many misunderstandings about depth perception.

Mr Howarth is right to mention the difficulties involved in investigations in indirect vision and I should like to say that I am not very satisfied with my results in monocular vision, which certainly cannot be regarded as 'significant'. More significant I hope are the results with double images.

May I be permitted to mention the fact, that you, Sir, know it was necessary to cut out a good deal of theoretical discussion as well as some experimental findings of my own and other people in order to meet the needs of the *British Journal of Psychology*. For this reason perhaps some points of my paper were not sufficiently clear.

I hope to publish some more results in the near future. These concern stereoscopic depth perception. I hope to base on them, as well as on previous investigations, one more theoretical approach to the interesting question of depth perception.

Yours faithfully

J. L. ZAJĄC

Department of Psychology
The University of Edinburgh
21 *July* 1953

PUBLICATIONS RECENTLY RECEIVED

Symposium on Fatigue. Edited by W. F. Floyd and A. T. Welford. London: Lewis and Co. Ltd. 1953. Pp. viii+196. 24*s.*

This volume includes most of the papers read at the 1952 Symposium of the Ergonomics Research Society. Since the Society publishes no journal, this public view of its activities is very welcome, and a discussion of fatigue is not only well suited to the Society's function of bringing together biologists of several disciplines, but also timely. The realization, after the first World War, that the problem of fatigue was less simple than it had at first been supposed, has been followed by a host of investigations thrust upon us by the problems of the second World War, and this is a good time to take stock of the progress made and the insights gained.

Bartlett, for example, in his opening paper, makes the good point that diminished output and increased effort, so often pointed to as criteria of fatigue, are not characteristics of performance as such, and that psychologists ought rather to look for criteria which themselves lie within the activities that are being studied. Papers by other Cambridge workers show what progress has been made in this direction.

Equally valuable to psychologists is the opportunity to take a look at current research within other disciplines. There is, for example, a good paper by Darcus from the anatomist's viewpoint, and Weston contributes a useful review of work on visual problems. Among several papers from Scandinavian authors is an interesting account, by Prof. Christensen of Stockholm, of the physiological valuation of work in Nykroppa Iron Works. It is to be hoped that the Society will continue to give English readers accounts of work not readily accessible to them.

The book, which contains twenty papers in all, has been very well produced, the quality of the photographs and diagrams meriting particular mention.

N. E. LOVELESS

Divine Horsemen. The Living Gods of Haiti. By Maya Deren. London and New York: Thames and Hudson. 1952. Pp. x+350. 25*s.*

This book is a difficult one to review adequately for the following reason. The writer states clearly that she is an artist and not a scientist and because of this it would be unfair to criticize her work from a technical standpoint. Her preface contains a claim which, in my opinion, should be given serious consideration by social psychologists and others working on deep analyses of cultural phenomena. The claim is that in investigations which involve non-verbal communication the artist is more at home than the academically trained social scientist. Miss Deren says: 'I am referring to the communication of concepts and ideas by means addressed to emotional and psychological perceptions rather than in terms of a self-conscious articulation or an address to intellectual analysis; for example the idea of power conveyed in posture rather than in the label of a name or definition.' And again: 'In effect, sensitivity to form provides the artist with a vast area of clues and data that might elude the professional anthropologist whose training emphasizes precisely that "scientific" detachment which may muffle even his normal sensitivity and responsiveness to formal nuance and subtlety, so that he becomes dependent upon the vagaries of informants' memory, intelligence and articulation.'

Miss Deren's book illustrates both the benefits and difficulties which can arise from the use of this method. On the positive side she has penetrated Voudoun to a depth which is rare in field work. On the other hand her genuine affection for the people and participation in their ceremonies has involved her to such an extent that she herself experienced possession. This involvement is perfectly understandable to anyone who has worked in the West Indies. It does, however, raise again the problem of how the investigator can participate enough to understand his people and yet retain that amount of aloofness which prevents him from being swept away by, for example, the insidious rhythm of drums.

This same involvement has made Miss Deren overemphasize the uniqueness of Voudoun as a religion. Voudoun appears to be a form of response to past and present social conditions which is prevalent with variations over the whole West Indies. The importance of her work to the scientist is that she has studied Voudoun in its most intensive form, where it is enough of an integrated institution to have a ritual which is explicable. In Jamaica anyhow the metaphysical system has become attenuated. In Haiti one is possessed by a specific loa, in Jamaica by 'the Spirit' or by that of a specific dead person. The difference is probably accounted for by the fact that Protestantism did not provide the necessary supernatural figures of the saints.

The language in which the book is written at times becomes somewhat esoteric. In spite of this there is much in it which is of the greatest value to those workers interested in the West Indies, in religion, in folklore and to those social psychologists who are interested in people rather than abstractions.

MADELINE KERR

The Origin of Intelligence in the Child. By JEAN PIAGET, translated by MARGARET COOK, London: Routledge and Kegan Paul. Pp. xi + 425. 25*s.*

The publication of an English edition of this important book will be generally welcomed by all psychologists concerned with the development of young children and by those concerned with the nature of intelligence and the formation of concepts. The book follows a very clear pattern, tracing the intellectual development of children from their first observed activities, an hour or so after birth until their behaviour can be said to be intelligent, somewhere about the age of eighteen months. The study is based on close observation of Prof. Piaget's own three children; there are 183 observations presented in considerable detail, some of them comparing the behaviour of all three, some tracing an activity over several days, all of them scrupulously giving the essential background of the observation and, from time to time, by reference back and forward, relating an incident to its antecedents or to subsequent behaviour to which it might be a fore-runner. Such a detailed set of systematic observations, by as skilled an observer as Piaget, is itself a valuable source of authentic material. This, however, is only the factual background for a very closely reasoned analysis of the nature of intelligence. The facts are collected systematically and presented so that there emerges a progressive pattern of development which is traced from birth through six successive stages. The starting-point is the schema of activity arising from the behaviour pattern produced by each of the primary reflex activities, sucking, grasping, etc. These schemata are extended by the assimilation into them of elements derived from the environment, the schemata being modified by accommodation to these elements. By mutual assimilation and accommodation, these schemata are grouped, through sensori-motor activity, and later regrouped to meet new situations until the final stage is reached, when some form of representation has been achieved, where invention of new means becomes possible by mental combinations instead of by active experimentation.

The final chapter of the book is devoted to a discussion of the relation of this system to the main theories of intelligence. It cannot be reconciled with the empirical-associationist or with the 'vitalistic-intellectualism' theories and, although it has some elements in common with gestalt psychology, a fundamental difference is found in the influence of acquired experience on the solution of new problems which is essential to Piaget's theory and superfluous in the evaluation of the pregnance of 'form'. After discussing pragmatic theories, he concludes with a lucid résumé of his own theory.

The book is stimulating and some of the interpretations brilliant, both in their clarity and in the way they illuminate the path to further developments. The whole is closely argued and makes a first reading far from easy. The inferences and deductions seem to follow so directly from the observations that one wonders that they have not been so expressed before. Now that his clear statement is available in English one can look forward to a period of criticism, analysis and further evaluation of the theory by crucial testing and observation. The book merits very serious study

F. W. LAND

The Postural Development of Infant Chimpanzees. By A. H. RIESEN and E. F. KINDER. New Haven: Yale University Press. London: Geoffrey Cumberlege, Oxford University Press. 1952. Pp. xxi + 204. 32*s.* 6*d.* net.

The study of early postural development in the chimpanzee described in this book is based on a series of examinations extending over the first year of life of a number of nursery-reared chimpanzees. The experimenters used the Gesell-Thompson schedule in an attempt to determine the comparative rates of development in the human and chimpanzee infant; in addition to these tests they used one, for estimating the chimpanzee's capacity to support its weight on a rod, which was not used with the human infants. Other aspects of development were assessed at the same time and the author promises future publication of the results. Although the official schedule examination started at 4 weeks, a few earlier observations are included because in the chimpanzee several functions appear and mature in these first weeks of life which only develop later in the human infant. Moreover, like the human child, the chimpanzee is born with or acquires responses which he later loses, a few of which he shows only during the first month of life.

The comparative findings may be summed up by saying that in all the grosser locomotor and postural skills the chimpanzee is considerably in advance of the human infant, while the latter develops finer movements in advance of, and probably to a greater extent than, the chimpanzee. Structural anatomical differences seem to account for differences in preferred posture and the manner in which muscular skills are most readily developed and used. The chimpanzees were more variable in their development than Gesell and Thompson's human subjects in spite of being reared in a much more uniform environment; this probably reflects the fact that the scale was designed for the description and study of human development and may be less truly representative for the normal development of the chimpanzee. Considerable differences in behaviour were observed between the human and chimpanzee subjects; the chimpanzee infants were more nervous and were fearful of unfamiliar objects and situations, and it seems possible that the highly restricted artificial conditions under which they were reared may have affected their emotional responses. In spite of the unnatural conditions of rearing, the authors report no significant differences in postural development between these nursery-reared chimpanzees and others which were reared by their mothers.

The book gives a detailed analysis of each item on the scale and a comparison for human and chimpanzee development on each item. The detail is considerable and at times a trifle confusing. Much of the material is printed in the form of tables which are not always easy to follow, the difficulty being increased by the fact that the legend is easily confused with the text. There are some excellent illustrative photographs.

One of the most interesting points suggested by the book (although the author does not discuss it) is that the precise psychological value of precocious postural development in the human infant may need to be reconsidered. The tendency among psychologists and nurse-maids to regard postural and locomotor development as correlated with intelligence appears a little unsatisfactory when one considers that the chimpanzee is more advanced in this respect than the human infant and the monkey even more precocious than the chimpanzee.

M. M. L.

The Human Senses. By F. A. Geldard. London: Chapman and Hall, Ltd. 1953. Pp. x+365. 40*s*. net.

Despite the somewhat excessive claim that an appreciation of man's senses is 'the highroad to the understanding of human nature', this is an orthodox and up-to-date survey of the physiological psychology of the sense organs. Its scope is defined by the traditional distinction of sensation from perception—a division which to the present reviewer seems sometimes to lead to rather arbitrary exclusions. For example, the basis of visual acuity is discussed, but that of the discrimination of visual depth is not, presumably because it is assigned to 'perception'.

Within these limits, the material is well presented. A notable feature is the generous treatment given to the senses other than vision and hearing. The chapters on the skin senses in particular are a considerable improvement on the thin fare usually provided in this field. It is a pity that the author did not find it possible to perform the same service for kinesthesis, which gets a bare seven pages.

N. E. LOVELESS

The Making of a Scientist. By Anne Roe. New York: Dodd, Mead and Company. 1953. Pp. xii+244. $3.75.

Written for a popular audience, this book reports an inquiry into the life history (elicited in interviews) and the mental characteristics (assessed by intelligence tests, the T.A.T. and the Rorschach) of sixty-four eminent American scientists, divided into three groups, physicists, biologists and social scientists (anthropologists and psychologists). It was evidently a serious and competently conducted inquiry, notable for the co-operation that Dr Roe won from her distinguished subjects. The results are presented here in the form of summarized impressions, with a minimum of supporting data but ample illustrative quotations from the interviews. A brief chapter on thought processes suggests the possibility that social scientists rely much more on verbal and much less on visual imagery than the other scientists. But the most interesting material comes from the life histories. There is suggestive evidence that the loss of a parent in childhood is commoner among these scientists than among the population at large, and Dr Roe points out the danger of generalizing too simply about the bad effects of a 'broken home'. She also found interesting differences in early experience between the social scientists and the other scientists, the former including more who had serious conflicts with their parents and more whose families thought themselves

superior to those around them; in later life they also included more who had been divorced. Irritating at times from its slightly kittenish style of writing, the book still deserves serious attention as reporting one of the very few psychological studies of eminent living people and as suggesting some interesting possibilities that would repay further inquiry.
D. W. H.

Analyzing and Predicting Juvenile Delinquency with the M.M.P.I. Edited by STARKE R. HATHAWAY and ELIO D. MONACHESI. University of Minnesota Press. 1953. Pp. viii + 153. 28*s*.

This book is concerned with the problem of identifying the predelinquent child through the evaluation of personality characteristics: it is a collection of papers which deal with the use of the M.M.P.I. for this purpose. In the first three inquiries reported, it was demonstrated that this test differentiates delinquent and non-delinquent adolescents. Further work was then aimed at exploring the efficiency of the test for predicting delinquency, the most important study being the follow-up of over 4000 schoolchildren to whom the inventory had been administered. Certain of the scales suggest that it might be possible to anticipate some delinquency by this method; the authors conclude, however, that it is easier to predict which children will not commit offences than it is to identify those who will. Although they claim valid results with early adolescents, it is doubtful whether younger children would be capable of completing the inventory satisfactorily; as a tool in the prevention of juvenile crime, therefore, it would miss the age of peak incidence. This makes the book more of academic interest than of practical value. Secondly, in suggesting a method of prevention which earmarks children who are considered prone to delinquency, the authors do not mention the danger of the label itself, through the attitudes of those applying it, acting as a precipitant.

The book is a useful collection of data on the subject, but the findings would be more readily grasped if the scales were referred to by name, rather than by letters or numbers, a practice which involves the reader in constant referring back.
M. W.

Statistics for the Social Sciences. By T. G. CONNOLLY and W. SLUCKIN. London: The Cleaver-Hume Press. 1953. Pp. vii + 154. 16*s*.

This book starts from the beginning and tells you just what you have to do with your raw material in order to produce the various statistics which are commonly included in presentations of experimental data in the fields of Psychology, Education and the Social Sciences. Virtually no previous statistical knowledge or experience is assumed. It embraces the usual techniques of calculating means, standard deviations, centiles, closeness of fit, normal distributions, '*t*' test of significance, the use of 'chi squared', sampling reliability, correlations, scatter diagrams, biserial and tetrachoric correlations, rank correlations and Spearman's foot-rule formula. Its main virtue would appear to be that it is a practical book from which actual methods can be applied directly to material available to the experimenter. The procedures to be adopted are clearly described and it should give the beginner a good start in actual practice. It would, I think, have been improved if it included some exercises and possibly an appendix containing some worked examples as patterns. It is a book for beginners which will enable them to acquire statistical experience, and it would seem to satisfy this limited purpose admirably. The chief defect is that although in a number of contexts the authors imply that there is, in fact, more to statistics than the mere calculation of a statistic, they give no hint of where a beginner should turn for more advanced guidance once the elements have been mastered. A selected bibliography would be helpful, as would even a suggestion that an experimenter might ask himself what use he intends making of his data before he begins to collect them rather than being content to tell him what he might do with them after he has gone to the trouble of collecting them. The book does fill a gap, and if a second edition is ever contemplated, the authors might bear in mind the needs of the more active readers who will want to carry their knowledge beyond the basic, uncritical, stage of routine calculation.
F. W. LAND

The Writing of Arabic Numerals. By G. G. NEILL WRIGHT. The Scottish Council for Research in Education. XXXIII. The University of London Press. 1952. Pp. xxii + 424. 10*s.* 6*d.*

As the author says, 'Arabic numerals are in everyday use....It is surprising to find how little has been published on the teaching of this indispensable skill.' The teacher who is looking for a manual of instruction, however, will not find this book wholly easy reading. In order to reach the valuable conclusions as to the best form of numerals for infant teaching and the best way to teach them, he must first digest a long though interesting section on the historical development of numerals and a second section on the author's own experiments on the writing of numerals by adults and children. The writer is obviously an enthusiast on his subject, however: although the reader may at first reel on learning that 'the numeral 1 can be written in more than eight hundred and fifty-seven million million million million ways', he soon finds himself studying with interest the numerous illustrations which show the development of numerals in a wide variety of different cultures and periods.

There are gaps in our present knowledge of how numerals have evolved. The author conceived the notion that through the study of how groups of people write numerals slowly and at speed some light might be thrown on the historical processes of development. To this end he conducted investigations with adult students, using among other methods that of serial reproduction. 'Both the historical and experimental data leave together the strong impression that every graphical sign tends to vary in a certain determinate direction.' In order to ascertain the best forms of numerals for infant teaching he carried out further experiments with schoolchildren. He gives some valuable facts on the numerals most likely to be written illegibly or interpreted ambiguously and suggests that some reconsideration should be given to the form of such numerals as, for example, 0, 7 and 2.

Readers interested in the historical development of numerals will find much to ponder over: the author's experiments will attract those concerned with the reproduction of visual material; educationists will welcome the last few chapters on the teaching of numerals in schools and the author's suggestions about method. But the volume suffers from the lack of a final chapter summarizing the work as a whole and its conclusions.

G.K.

Visibility. A bibliography. Compiled by JACK WEINER and MORRIS C. LEIKIND. Edited by JACK R. GIBSON. Washington: The Library of Congress. 1952. Pp. vi + 90. 65 cents.

This bibliography of over 2000 items has sections dealing with laboratory investigations, factors such as atmospheric conditions affecting visibility in natural circumstances (including a section on underwater photometry), methods of measurement in the field, and 'engineering' applications of visibility data, including camouflage, blackout, navigation aids and signalling.

PROFESSOR CLARK L. HULL

VOLUME XLV MAY 1954 PART 2

CRITICAL NOTICE AND APPRECIATION OF THE WORK OF THE LATE PROFESSOR CLARK L. HULL

BY F. V. SMITH

Department of Psychology, University of Durham

A Behavior System. BY CLARK L. HULL. Yale University Press. 1953. (London: Geoffrey Cumberlege. English price 40*s.* net.)

The peculiar difficulties of explanation in Psychology are indicated in the many and diverse systems which have appeared from time to time and particularly in the present century. Nearly all of them have enhanced our understanding of at least some facets of behaviour and some have proved more susceptible of extension and application than others. It could be claimed that the Behaviouristic method did not begin with J. B. Watson's forthright pronouncements; but despite apparent departures from or weaknesses in planning and consequent inconsistencies in terminology, Watson had suggested a broad plan which others were to follow and the most notable of these, as builders of systems, have been Prof. E. C. Tolman and the late Prof. Clark L. Hull. Tolman's *Purposive Behavior in Animals and Men* (1932) presented an ingenious system based upon the two assumptions that (*a*) 'Behaviour, as behaviour, that is, as molar, is purposive and is cognitive. These purposes and cognitions are its immediate descriptive "warp and woof"', and (*b*) that, 'such purposes and cognitions are just as evident, as we shall see later, if this behaviour be that of a rat as if it be that of a human being'. Tolman's attribution to the environment of features of a kind which are no doubt significant for the animals, like Lewin's allocation of degrees of significance reflecting the condition of the organism, gave an interesting review of the possibilities within a particular behavioural situation at a given time; but as Lewin himself conceded, such allocation presumed that all the problems of Psychology had been solved, and it is impracticable, except in repeated trials, to determine with any precision the several 'valences', 'topological stresses', 'discriminanda', 'manipulanda' and 'manipulanda expectations' and the relationships between the kaleidoscopic changes from one moment to another.

Hull's intention was to begin in a different way. As he wrote (1943*a*), 'an ideally adequate theory even of so-called purposive behaviour ought, therefore, to begin with colourless movement and mere receptor impulses as such, and from these build up step by step both adaptive behaviour and maladaptive behaviour. The present approach does not deny the molar reality of purposive acts (as opposed to movement), of intelligence, of insight, of goals, of intents, of striving, or of value; on the contrary, we insist upon the genuineness of these forms of behaviour. We hope ultimately to show the logical right to the use of such concepts by deducing them as secondary principles from more elementary objective primary principles'. Throughout the development of such a system, the criteria of explanation were to be that, 'an observed event is said to be explained when the proposition expressing it has been logically derived from a set of definitions and postulates coupled with observed conditions antecedent to the event' (1943*b*). The science of behaviour was regarded as 'a true, though immature natural science'. (Personal communication, 20 Oct. 1949.)

In many ways the *Principles of Behavior* (1943) offered an integration of the treatment of facets of behaviour which had appeared in journals since the late 'twenties. Notable among these were: 'Knowledge and purpose as habit mechanisms' (1930), 'The goal-gradient hypothesis and maze learning' (1932), 'The mechanism of the assembly of behaviour segments in novel combinations suitable for problem solution' (1935) and 'Mind, mechanism and adaptive behaviour' (1937), the latter being an account of a miniature system in which aspects of purposive behaviour were deduced in thirteen theorems from a number of simple postulates. The interesting feature in these papers was that a behaviouristic explanation was attempted for 'insight', 'antedating reactions' and some types of problem-solving, phenomena hitherto 'explained' by the mention of the term 'insight'—and the process of understanding for the reader began with simple empirical findings associated with conditioning.

It is obvious that the extent of such systems, as represented by the number of postulates and definitions, the deductions therefrom and possible applications is limited by our knowledge of the functions of nervous systems and reliable observations of behaviour; but the *Principles of Behavior* did present the most rigorous and extensive behaviouristic system to date and clearly involved considerable erudition and painstaking care in collating a large number of empirical findings. There were sixteen postulates and a number of corollaries which dealt with several factors relevant to the adjustment of the organism to its environment. These included Reinforcement, Stimulus Generalization, Motivation, Inhibition, Oscillation and Reaction Latency. The basic concept which provided the instigation for the system was that of need, which was claimed to exist 'when a condition arises for which action on the part of the organism is a pre-requisite to the optimum probability of survival of either the individual or the species' (1943*c*). The need was claimed to produce the primary animal drives. One could ask whether or not Hull had contradicted his plan to refrain from introducing such concepts as purposive action, insight, goals, intents and striving until they could be deduced 'as secondary principles from the more elementary objective primary principles'. The impression that many of the quantitative examples were artificially simplified and that in the absence of conclusive data or indications as to method, the model of the physical sciences was invariably imposed, did arise; but detailed queries and points of disagreement should not be allowed to detract from the achievement as a whole. Hull had drawn attention to the possible inter-relation of a number of concepts involved in adaptive behaviour. The general plan admitted of extensions and refinements (which Hull and his co-workers pursued assiduously and published frequently in cyclostyled form) and of the behaviour covered by the system, the reader did feel that he had gained an understanding not hitherto available to him.

The continued work of refinement found expression in the *Essentials of Behavior* published in 1951. Once more the basic concept was need reduction and the eighteen postulates described again the fundamental processes by which behavioural adjustments could be made and further adaptations deduced. The order of the postulates is slightly different from that in the *Principles of Behavior* and the wording and mathematical expression has in several places been revised. A criticism which was anticipated in the *Principles of Behavior*, and discussed elsewhere by Hull (1945), is met directly by the last postulate—No. XVIII—'The "constant" numerical values appearing in equations representing primary molar behavioural laws vary from species to species, from individual to individual, and from some physiological states to others in the same individual at different times, all quite apart from the factor of behavioural oscillation ($_{S}O_{R}$)'.

In a personal communication dated 20 October 1949, Hull mentioned a further work 'on individual, as distinguished from social, behaviour' or as he describes it in the preface to *A Behavior System* published in 1952, 'intended primarily to show the application of the *principles*[1] to the deduction of the simpler phenomena characterizing the behaviour of single organisms'. There was too the plan for yet another work which 'would apply these same principles—to the deduction of the elementary phenomena of social behaviour, i.e. of behaviour manifested when the interacting objects are mammalian organisms of the same species'; but in the same preface Hull states, 'I greatly regret that in all probability I shall not be able to write the third volume'. In his letter of 20 October 1949, Hull indicated that he 'was very ill with coronary thrombosis....This condition combined with the fact that I am sixty-five years of age, greatly limits my capacity for work.... If I live to the age of retirement (68 at Yale) I ought to finish the volume....' As Prof. Mark A. May points out in a Foreword to *A Behavior System*, 'Hull won the race'. He died on 10 May 1952, after the MS. had gone to press, and did not see the galley proofs.

The book is admirably arranged and printed. There are eighteen postulates and fifteen corollaries in *A Behavior System*, following broadly the same lines as in the *Essentials of Behavior*. These are set out at the beginning with the mathematical equations for the postulates and their subsections numbered consecutively, giving to date, thirty-one basic equations, many of which embody provisional numerical approximations. The postulates may be regarded as the most defensible form of statement available to the author in regard to a particular aspect of behaviour, such as the constitution of Reaction Potential, Inhibitory Potential or Stimulus Generalization. They are used in the latest work more as summarizing a general background of knowledge upon which the processes of reasoning in the 133 theorems are based. Theorem 1 deals with the course of the development of two super-threshold, reaction potentials in competition. Theorem 30 deals with the results of shifts in incentives as suggested by the work of Crespi (1942)—'Other things constant, an abrupt shift in the incentive used during a maze-learning process will be followed first by a major shift in reaction potential and then by two or more progressively smaller shifts on successive trials, the series constituting a rapid learning process of the exponential variety, culminating in the course that $_{S}E_{R}$ would have followed had the new incentive been operating continuously from the beginning of the learning'. Theorem 48 states, 'other things equal, the repeated performance and reinforcement of simple behaviour links tend to shift the contraction-intensity phases of the several muscles in such a way as to attain the maximum of reinforcement which is consistent with a minimum of work'. Theorem 96 states: 'When an indirect member of a locomotor habit-family hierarchy has attained a goal in a novel situation involving relatively free space, the learning then acquired is transferred to the initial segments of remaining members of the hierarchy without specific practice and on subsequent trials is manifested by the organism's spontaneous choice of the most direct path'.

Perhaps the most interesting extension of Hull's approach for those working in the social sciences is the clearly defensible attempt in chapter eleven to treat the concept of value and valuation behaviouristically. The treatment begins with a consideration of nutritional needs which have an interesting and methodologically convenient evolutionary continuity, and while the scheme is of necessity not worked out in detail, enough is indicated to suggest that an extension to other activities is possible.

---

[1] Italics mine.

In a personal communication dated 29 February 1952, Hull, with characteristic modesty, wrote, 'when one is in the midst of system making the great gaps in the structure are so evident that the task is very discouraging'. Few realized more clearly than Hull the limitations of his systems. Of *A Behavior System*, he could write in the same letter, 'More than ever I realize the inadequacy of the work—the parts I have left out. Really any system written at present must be a miniature one in comparison with the true system as it will some day be known'; but the estimate of Hull as an architect of systems cannot be influenced by considerations of extent alone, nor the very provisional numerical coefficients and the artificially simplified examples. He was intent upon indicating lines of development and was ever ready to revise when further information came to hand, regarding agreement of the deductions made with empirical fact as the criterion of the soundness of the system. To arrive at the same result by two entirely different routes, that is by a process of deduction from basic axioms or postulates and by a process of generalization from empirical findings is certainly an effective method of allaying the provocation of a human mind intent upon explanation and, despite the inevitable limitations, in his time, Hull's systems do leave the student with the impression of penetration and the prospect of further extension as and when reliable behaviouristic designations of relevant states become possible. It is of interest to note that 52% of the 178 theoretical propositions in his latest work were thought to have empirical confirmation and a further 17% to have approximate or indirect evidence in support.

It would be easier in 1970 to assess Hull's influence upon the course of Psychology; but without undue hazard, attention could be drawn to certain features in 1954. There is the definite stress on objectivity and experimental method, apparent in his *Aptitude Testing* (1928, 1929) and *Hypnosis and Suggestibility* (1933) which marked a departure from much of the writing in the earlier years of this century when reference to experimental test lagged far behind the invention of plausible and attractive hypothetical systems. There is the improved orientation in the fields of general behaviour and learning which derives from his efforts in collating the experimental evidence relating to reinforcement, stimulus generalization, inhibition, oscillation and reaction latency, the consequent stimulus to investigation in these fields and, perhaps most important of all, there is the acceptance of the challenge that many of the questions addressed to Psychology are in the form of those addressed to a natural science, indicating the desire to know how the particular behavioural events take place. In response to this challenge there are the works, *Mathematico-deductive Theory of Rote Learning* (Hull, Hovland, Ross, Hall, Perkins & Fitch, 1940), *Principles of Behavior*, *Essentials of Behavior* and *A Behavior System*, in his own view as yet sadly incomplete as systems of explanation but, in principle, capable of extension not only in the fields in which he and his closer colleagues were engaged, but to such topics as attention, see Berlyne (1951), the vexed literature of emotions, see Brown & Farber (1951) and in J. G. Taylor's forthcoming work on Perception.

In conclusion, Hull's personal achievement, in bringing to a definite point his contribution to the efforts of the Institute of Human Relations at Yale to develop a basic science of behaviour, deserves special mention. The foregoing is essentially an academic appreciation of Hull, since we never met. I had hoped to meet him at the International Congress of Psychology in Edinburgh in 1948; but he was too ill to come. It is more appropriate that the appreciation of his personal qualities should be made by those who worked with him for many years; but one feels bound to acknowledge a prompt and generous correspondence, always marked by scientific humility and appreciation of help from his colleagues.

REFERENCES

BERLYNE, D. E. (1951). Attention, perception and behaviour theory. *Psychol. Rev.* LVIII, 137–54.

BROWN, J. S. & FARBER, I. E. (1951). Emotions conceptualised as intervening variables—with suggestions towards a theory of frustration. *Psychol. Bull.* XLVIII, 465–95.

CRESPI, L. P. (1942). Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.* LV, 467–517.

HULL, C. L. (1928, 1929). *Aptitude Testing*. Lond.: Geo. Harrap & Co.

HULL, C. L. (1930). Knowledge and purpose as habit mechanisms. *Psychol. Rev.* XXXVII, 511–25.

HULL, C. L. (1932). The goal-gradient hypothesis and maze-learning. *Psychol. Rev.* XXXIX, 25–43.

HULL, C. L. (1933). *Hypnosis and Suggestibility*. New York: Appleton-Century-Crofts, Inc.

HULL, C. L. (1935). The mechanism of the assembly of behavior segments in novel combinations suitable for problem solution. *Psychol. Rev.* XLII, 219–45.

HULL, C. L. (1937). Mind, mechanism and adaptive behaviour. *Psychol. Rev.* XLIV, 1–32.

HULL, C. L. (1943). *Principles of Behavior*. (*a*) pp. 25–26; (*b*) pp. 2–3; (*c*) p. 57. New York: Appleton-Century-Crofts, Inc.

HULL, C. L. (1945). The place of innate individual and species differences in a natural-science theory of behavior. *Psychol. Rev.* LII, 55–60.

HULL, C. L. (1951). *Essentials of Behavior*. New Haven, Conn.: Yale University Press.

HULL, C. L. (1952). *A Behavior System*. New Haven, Conn.: Yale University Press; London: Geoffrey Cumberlege.

HULL, C. L., HOVLAND, C. I., ROSS, R. T., HALL, M., PERKINS, D. T., and FITCH, F. B. (1940). *Mathematico-deductive Theory of Rote Learning*. New Haven, Conn.: Yale University Press.

THE EFFECT OF THE VOCATIONAL AIMS OF INDUSTRIAL APPRENTICES UPON THEIR ATTITUDE TO EDUCATION AND COMMISSIONING[1]

By SYLVIA ANTHONY

Air Ministry, London

I. *Introduction* (p. 82). II. *The population* (pp. 82–83). III. *Method* (p. 83). IV. *Rejection of educational opportunity* (pp. 83–85). V. *Rejection of commissioning opportunity* (pp. 85–86). VI. *Discussion* (p. 87).

I. INTRODUCTION

The aim of this paper is to show that the vocational aspirations of a large section of the population of the United Kingdom lead them to reject opportunities of filling roles associated with superior social status. First, this section of the population will be defined; secondly, I shall attempt to show that many included in this group have opportunities for filling superior status roles; and thirdly, that they tend to reject such opportunities. Lastly, the motives of the behaviour described will be discussed.

II. THE POPULATION

The population studied consists of a sample of industrial apprentices; men who spend five years between ages 16 and 21 learning a skilled trade. Apprentices enter the R.A.F. from about 400 skilled trades (that is to say, the names of distinctive occupations listed for apprentices entering over a period of 1½ years totalled about 400), the range and quality of which may be briefly illustrated by quoting twenty from an alphabetically ordered list: *Radio mechanic, radiographer, railway engineer, reed maker, reel maker, refrigerator mechanic, rivetter, rockman, roll turner, ropemaker, rubber worker, saddler, sailmaker, sanitary engineer, saw doctor, sawsmith, sawyer, scale fitter, scientific instrument maker, scythe grinder.*

The training given and the skill acquired by the apprentice vary greatly in different trades, and even within trades between different areas and employers. These sources of difference do not now concern us. The educational level of apprentices also varies greatly. Many an apprentice continues his formal education regularly at evening classes and perhaps also at day continuation school, taking examinations from the preliminary level through the national engineering series, sometimes up to B.Sc. Others discontinue formal education a year or two before beginning apprenticeship and may continue to neglect it. The range of educational achievement descends to the borderline of illiteracy.

Apprentices form about one-quarter of the whole population of young men accepted for national military service annually in the United Kingdom, as will be seen from the analysis in Table 1.

The majority of national service recruits are called up at about 18 years of age, but apprentices usually at the age of 21. In a representative sample of 2% of all national service recruits entering the R.A.F. in the year following 1 August 1951, 66% entered at the age of 18, about 8% after deferment for University or other professional studies, and about 25% after deferment for completion of apprenticeship. This sample of apprentices numbered 224 men, and represented over 11,000 apprentices entering the R.A.F. during the year.

---

[1] Paper delivered at a meeting of the British Association for the Advancement of Science, Liverpool, September 1953. Published by permission of the Air Ministry.

III. METHOD

The grammar school is an avenue to superior social status. The Queen's commission confers such status. To show that industrial apprentices tend to reject opportunities for filling superior status roles, it must be shown (i) that a certain proportion of them fulfil the elementary qualifications for admission, (ii) that a proportion of those so qualified reject such opportunities, and (iii) that the proportion so behaving is significantly greater than that in a group representative of a non-apprentice population under comparable conditions. The method employed was (*a*) to compare the scores on a test of general intelligence and the schooling of the apprentices with those of the main body of the 2% recruit sample, taking into account also the statements of individuals on their educational experience in relation to their vocational aims, and (*b*) to compare the proportion of applications for commissions made by apprentices with those of other national service recruits under similar conditions, using another sample comprising only recruits eligible to apply for commissions.

Table 1. *Disposal of men liable for military service under the National Service Act* (1948)

(Average numbers for age classes of 1933–1934 inclusive.)

Disposal	Registered No.	Registered %	Accepted (%)
Posted to H.M. Forces (as N.S. entrants or as volunteers)	144,700	51·5	66·2
Deferring			
(*a*) for apprenticeships	53,500	19·1	24·5
(*b*) for University or professional studies	20,400	7·3	9·3
Total acceptable	218,000	77·9	100·0
Exempt owing to occupation	22,900	8·1	—
Unfit	36,900	13·2	—
Others	2,100	0·8	—
Total registering	280,500	100·0	—

Entries calculated from statistics published in the *Ministry of Labour Gazette*, May 1953.

The two opportunities studied do not occur entirely independently. The preliminary screening of national service candidates for commissions is by an educational qualification (formerly the possession of a School Certificate or a certificate of equivalent standing). Of the total national service recruit sample studied, 37·5% qualified in this way. Of these qualified men, 80% had attended grammar or public schools and 14% technical secondary schools full time. Entry to technical schools is frequently determined mainly by results on the examination which governs grammar school entry.

IV. REJECTION OF EDUCATIONAL OPPORTUNITY

On entry every recruit is given certain tests of general intelligence and aptitude. He also fills in a form giving particulars of his education, including the school or college last attended full time. Details are later amplified and checked in individual interviews. At the interview, some men who have entered *secondary modern* for their last full-time schooling, refer to the examination at the age of 11+ and say that they voluntarily did not sit, or that they had to enter but did not wish to succeed in securing grammar school education.

These verbal responses, recorded by the interviewers, have not been systematically studied because the interviews are not rigidly standardized, and current report of attitudes held seven to ten years previously is not reliable evidence of the motives of behaviour at that time. But other evidence of the relation between educational background, abilities and apprenticeship of recruits suggests that such statements may represent not a sour grapes attitude nor any form of defence against a sense of inferiority, but a constant negative reaction to opportunities for grammar school education. The evidence is presented in tables of very simple form. The schooling statistics conceal regional and the mean test scores individual variation, which have both, however, been studied in detail. Analyses of variance support the argument presented by simple comparison of means and percentages.

Table 2. *Apprenticeship and post-primary education*

Post-primary education	National service recruits to the R.A.F.			
	Apprentices	Others	Total sample	
			%	No.
Secondary modern	57	40	45	376
Grammar, public or technical full time	43	60	55	466
Total, %	100	100	100	—
Total, No.	224	618	—	842

Table 3. *Apprenticeship and abilities*

R.A.F. test	Mean scores			Score range
	Apprentices	Others	Total	
General intelligence	37·46	37·74	37·67	0–60
Verbal comprehension	51·24	51·49	51·42	0–80

It will be seen from Table 2 that the apprentice entry includes a significantly larger percentage of men whose last full-time schooling was secondary modern or equivalent. (By χ^2 test of significance of difference, $P < 0{\cdot}001$.) Yet the average score made by these apprentices on two paper-pencil tests of ability is equal to that made by other recruits. The fractional differences recorded in Table 3 are not statistically significant, by t test.

It is suggested that the divergence in schooling between the two groups arose because the apprentice group included boys whose intelligence was associated, level for level, with less wish to succeed in the 11+ examination for grammar school places. The findings would also be consistent with the exclusion of only very dull boys from apprenticeship, in an otherwise representative sample. The latter alternative may be dismissed, since apprentice scores range as low as those of the other group and the total variance of scores of the two groups shows no significant difference (by χ^2, P approximates 0·3).

The R.A.F. tests aim to measure general intelligence and verbal comprehension. The 11+ examination for entry to grammar school aimed at the same sort of thing with the same men some years earlier. Table 4 offers some evidence that the two procedures were assessing similar or closely associated attributes. The average score of the secondary modern old-boy on R.A.F. tests is significantly lower than that of the boy who succeeded in the 11+ examination. The findings are suggestive, not conclusive, because in these data the effect of different secondary schooling upon R.A.F. test results is not separately measurable. By t test, $P < 0{\cdot}001$ that the differences between column entries in Table 4 would occur by the chances of random selection.

It appears then that the apprentice group is on average as intelligent as the rest of the body of R.A.F. recruits, but its success in the grammar school entry examinations is less than average, which may be because in many cases success would have conflicted with the aim of training to be a highly skilled tradesman.

V. REJECTION OF COMMISSIONING OPPORTUNITY

These decisions about schooling were made years back; the decisions about commissioning are being made to-day. Every Air Force national service recruit with the required educational qualifications is encouraged to apply for a commission for the period of his national service. The number of commissions granted is small compared with the number of men qualified to apply, and may vary periodically according to manning requirements. But within any one period the selectors' aim is to keep the standard constant.

Table 4. *Post-primary education and abilities*

(2 % sample of national service recruits to the R.A.F. over a year.)

	R.A.F. test mean scores		
Post-primary education	General intelligence	Verbal comprehension	No. of men
Secondary modern	33·24	44·35	376
Grammar, public or technical (full time)	41·23	56·89	466
Total sample	37·67	51·42	842
Test score range	0–60	0–80	—

Table 5. *Apprenticeship and commissioning applications*

	Recruits educationally eligible		Total	
	Apprentices (%)	Others (%)	(%)	No.
Apply for commissions	67	81	78	870
Do not apply	33	19	22	243
Total, %	100	100	100	—
Total, No.	—	—	—	1113

Those educationally qualified are called Potential Officer Material—P.O.M. Out of over 11,000 apprentices entering during the year, over 2500 were P.O.M. These men go to a camp together with the other P.O.M. for basic recruit training, and while there they are all invited to apply for commissions, interviewed by selecting officers, and if considered suitable, sent forward for the final decision of an Air Ministry Selection Board. (In addition to the national service P.O.M. from this camp, the Board also interviews a number of University graduates before their call up, who, if accepted, proceed direct to officer cadet training or to a medical officers' unit. Consequently the figures tabulated are not a complete account of the commissioning of national service recruits.)

Using the documentary returns from the P.O.M. camp over a period of four consecutive weeks, we may compare the percentage of apprentices and of non-apprentices who applied for commissions, and the percentage of these applicants who were passed by the preliminary and final selectors. It is clear from Table 5 that the percentage of apprentices applying for commissions was much lower than that of the other eligible recruits. The difference between the column entries is highly significant statistically ($P<0{\cdot}001$, by χ^2 test).

Clearly the apprentices have imposed upon themselves a severe self-screening process by which, on average, one in every three men eligible has decided *not* to apply for a commission. The self-screening of the other recruits is less severe—four out of five apply.

Following the applicants through, we find that selection on the station gives the results shown in Table 6.

Again the differences between apprentices and others is significantly greater than would be likely to arise by chances of sampling (by χ^2 test, $P<0{\cdot}001$). The selectors on the camp send forward a higher proportion of the apprentice applicants, 35% as against 11%, and so tend to correct the men's own severe self-screening. This finding was unexpected both by the researcher and by the selectors. The selectors were anxious that all eligible recruits should volunteer for commissioning and be given equal consideration, but they had not looked upon the apprentices as better candidates than the others (which, counting all those eligible, they may not have been), and had therefore not considered the apprentice *applicants* as better candidates than the others. But the selectors' actual assessments of individual cases had remained independent of their general judgements.

Table 6. *Apprenticeship and commissioning: provisional selection*

	P.O.M. applicants for commissions			
	Apprentices (%)	Others (%)	Total over 4-week period	
			(%)	No.
Recommended	35	11	15	131
Rejected	65	89	85	739
Total applying, %	100	100	100	—
Total applying, No.	—	—	—	870

Table 7. *Apprenticeship and commissioning: final selection for O.C.T.U.*

	Recommended by provisional selectors			
Final selection for training	Apprentices (%)	Others (%)	Total	
			(%)	No.
Selected	23	29	27	35
Rejected	77	71	73	96
Total recommended, %	100	100	100	—
Total recommended, No.	—	—	—	131

In the final selection for officer training, as shown in Table 7, the percentage of apprentices recommended by the station selecting officers and passed by the Board is somewhat lower than that of the other recruits, but the difference is statistically insignificant. That is, the Air Ministry Selection Board tends to pass in about the same proportion the apprentices and the other recruits sent forward by the selection officers who undertake the preliminary selection screening on the camp.

This concludes the evidence presented to establish the proposition that industrial apprentices tend to care less than other members of the population about acquiring officer rank and grammar school education. But it tells little about the impulses behind this behaviour. An attempt is in progress to find out more, by interviews and questionnaires given to apprentices on the P.O.M. training unit, and although the results are not complete, they give rise to suggestions that may be of use in interpreting the total material.

VI. DISCUSSION

We appear to be concerned, then, with two forms of social impulse, not primary, but operating in adults with some independence; a wish to acquire and practice an occupational art or skill, and a wish to acquire or maintain high social status. These impulses may act in parallel, and questions of their relative strength then tend to be overlooked. Alternatively, in certain circumstances, they may co-exist at cross-purposes. The way a man acts will then tend to be related to the degree of divergence and the relative strength of these impulses, and his opportunities for realizing them in action.

Two groups have been studied, in two fields of action: the apprentices and the remaining body of recruits, as applicants for secondary school places and for service commissions. In both fields it is generally agreed that considerations of social status operate as an incentive to achievement. In both the apprentice group achieves less than its abilities warrant. The vocational aims of the apprentices are unusually strong and clearly defined. They are not directed towards roles bearing high social status. It appears probable that the observed short fall in achievement is largely due to the divergence of the two aims and the greater strength of the vocational impulse.

Whether the desire for social status is weaker among the apprentices than among the general body of recruits does not appear. There is some evidence in the more recent data that temporary frustration of the vocational impulse permits a desire for social status to operate as strongly in the apprentice as in the general group and perhaps often more naïvely. In the officer selection process the total apprentice group tends to divide into two, one including those who have a hope of getting their own trade in the Air Force, and the other, those who have not. The attitude to officer rank of what we may call the trade-frustrated group, is different from that of the apprentice group in general. The trade-frustrated group rate the opportunity to be commissioned more highly.

Finally, a brief reference to social values. Because desire for higher social status frequently acts in line with vocational impulse, the distinction is apt to be overlooked. It may even be suggested that the two impulses *ought* to act in parallel—every man *should* desire grammar school education and commissioned rank. Actually the social health of the community may be positively associated with the strength of vocational impulses rather than with desires for higher social status. It would follow that the attitudes of the apprentice-artisan group, as they have been described, are not to be disapproved or discouraged.

(*Manuscript received* 18 *October* 1953)

THE EFFECT OF REFERENCE GROUP FUNCTIONS ON SOCIAL STATUS RATINGS[1]

By GURTH HIGGIN

Air Ministry London

I. *Introduction* (p. 88). II. *Method and source of data* (pp. 88–89). III. *Hypotheses formed on initial data* (pp. 89–90). IV. *Method of checking hypothesis* (p. 90). V. *Results obtained* (pp. 90–92). VI. *Summary* (pp. 92–93). *References* (p. 93).

I. INTRODUCTION

This paper is designed to show how a person's rating of the status of his social role can be affected by the satisfaction he gets from that role. For example, if being an Air Force sergeant is a very satisfying and important thing in a man's life, then when he is asked to rate the social status of the role of Air Force sergeant, he will tend to rate it higher than most people would. In the same way a man from a lorry driver's family who has become an Air Force officer will probably want to reject his association with his social background and will therefore rate its status lower than would most other people for whom it has no such negative associations. A group to which a person belongs and membership of which has some such marked value for him I shall call a 'reference group' (Hyman, 1942; Newcomb, 1943). This is only one usage of the term reference group which, in other usages, may not be restricted to membership groups (Kelley, 1952). The material to be drawn on is taken from a larger study which was not designed to demonstrate this hypothesis, so support for it can be given in general terms only and no detailed study of the social and psychological processes involved is attempted.

II. METHOD AND SOURCE OF DATA

The work from which the material was taken was concerned with the social mobility experienced by members of the R.A.F. in moving from civilian life to Air Force life. The two points of reference used were the social status of a man's family as represented by his father's occupation and the social status of his Air Force role. From each man a subjective and an objective rating of the social status of both these roles was obtained. Both types of rating were made in terms of seven typical occupations forming a standard seven-point scale adapted from the one set up by Hall & Jones (1950). The subjective ratings were the man's own assessments of the positions on the standard scale at which the two roles belonged. The objective ratings of family backgrounds were made on the basis of the Hall-Jones criteria, while the objective ratings of Air Force roles were arrived at from this research.

The seven-point scale of occupations used as a standard ranged from professional to unskilled manual occupations through the following groupings. The example of each given in parentheses is the example used in the research. Group one, professional and higher administrative (company director); group two, managerial and executive (works manager); group three, higher grade inspectorial and supervisory (elementary school teacher); group four, lower grade inspectorial and supervisory (head clerk); group five, routine non-manual and skilled manual (clerk, carpenter);[1] group six, semi-skilled manual (lorry driver); and group seven unskilled manual (dock labourer).

---

[1] Paper delivered at a meeting of the British Association for the Advancement of Science, Liverpool, September 1953. Published by permission of the Air Ministry.

The sample used was a group of 194 R.A.F. personnel made up of random samples of airmen (both regular and National Service airmen), non-commissioned officers, and officers. It was found that the mean rating given the standard occupations by the men in all these groups was the same for each of the occupations forming the standard scale as these occupations were given in the original Hall-Jones Study. From this evidence of the reliability of the use of the standard scale by the sample, it was assumed that the mean ratings given to the Service roles by the whole sample could be taken as a sufficiently accurate assessment of their social status in terms of the standard scale. These ratings were therefore taken as the objective status ratings of the Air Force roles. There were four roles involved—Regular Airman, National Service airman, R.A.F. Senior N.C.O. and R.A.F. Officer, one or other of which of course applied to every man in the sample.

With these four service ratings it was possible to construct an objective and a subjective mobility rating for every man. These mobility ratings were both expressed in terms of distance along the standard scale. The objective mobility rating was the number of steps between the objective status ratings of a man's family and of his R.A.F. role. The subjective rating was the scale distance between the man's own assessments of these two roles.

III. HYPOTHESES FORMED ON INITIAL DATA

Comparison of objective and subjective mobility ratings disclosed that in many cases these were different. Many of the minor differences could be accounted for from the nature of the rating technique which had been used. In using the technique it was possible for subjects to make a rating between points on the standard scale as well as on a level with them. However, even when minor differences arising from this cause had been allowed for, there was still quite a number of men who had either overestimated or underestimated their social mobility in joining the Air Force. It was, of course, possible that these discrepancies could have arisen from inability to rate occupations and R.A.F. roles accurately. In rating the standard scale, however, it had been found that all groups in the sample had used the status rating scale with considerable accuracy. It was therefore unlikely that inability to use the scale, or a failure to understand it, would explain the discrepancies observed in the assessments of mobility.

An hypothesis was developed to explain this deviant status rating behaviour in terms of reference group functions. It was suggested that if a man's R.A.F. role or his civilian social status had a strong reference group association for him, he would tend to rate such a group abnormally because of this. This supposition was based on the fact that ratings of social status as well as being concerned with recognizable facts about the prestige differences of social groups, also tend to be value judgements. In general, high status is considered to be more desirable than low status. Consequently, when rating a group with which one likes to be identified (that is, a positive reference group), there is a tendency to raise its status, Consequently, when rating a group of which one is a member, but with which one does not like to be identified (that is, a negative reference group) one tends to lower its status. Now the present hypothesis is that many men for whom their R.A.F. roles represent a positive reference group will tend to make abnormal ratings of their social mobility in joining the R.A.F. because they will overrate their civilian status. As a corollary to this it would also be expected that men who have adopted an R.A.F. reference group might depress their native civilian status group because (especially for the upwardly mobile) it may have acquired a negative reference for them. Similarly, men whose principal reference group is civilian might underrate their R.A.F. group as one with which they have no particular desire to be associated.

---

[1] Group five was originally divided into 5A (routine non-manual), and 5B (skilled manual). But as only 2% of the sample fell in group 5A, it was combined with 5B as one group.

IV. METHOD OF CHECKING HYPOTHESIS

In order to analyse the ratings of the men in the sample along these lines, the men were divided into three groups, according to the type of mobility ratings they made. These groups are labelled positive mobility raters, neutral mobility raters and negative mobility raters. The positive mobility raters were either men who had been upwardly mobile in joining the R.A.F., and who overrated the extent of this upward mobility, or men who had been downwardly mobile in joining the R.A.F. and who had underrated this downward mobility. The neutral mobility raters were those whose assessment of their mobility agreed with or was not greatly different from an objective rating of it. The negative mobility raters were either men who had been upwardly mobile and who underrated this upward mobility, or men who had been downwardly mobile and who overrated this downward mobility.

If the hypothesis just outlined is to be substantiated, it is necessary to find some evidence that the positive mobility raters take an R.A.F. role as a positive reference group significantly more frequently than do the negative mobility raters. And on the other hand, evidence is needed to show that the negative mobility raters look to a civilian reference group more often than the positive mobility raters. As has already been said, this small study is a by-product of a larger research which was not designed to test this particular hypothesis. The only direct evidence is in the status rating behaviour of the two groups, and because of small numbers this cannot be analysed in any great detail. There are, however, several lines of indirect evidence.

V. RESULTS OBTAINED

From the research a rating was available for each subject as to the relative satisfaction or dissatisfaction he felt towards his life in the R.A.F. In Table 1 the reactions of the two groups who made abnormal ratings of their civilian to service mobility are shown.

From Table 1 it can be seen that the positive mobility raters are significantly more often satisfied with R.A.F. life than are the negative mobility raters, while the latter group contains a significantly greater proportion of men who express dissatisfaction with R.A.F. life. These differences are in the direction predicted by our hypothesis and so provide some support for it.

If the process indicated in our hypothesis were the only reason why men tend to assess their mobility incorrectly we would expect the differences in the degree of satisfaction expressed by the two groups to be rather larger than they are. There are probably two reasons for this. First, in fact our hypothesis is not the only explanation of the tendency to make abnormal ratings of mobility, and secondly, the reason might lie partly in the shortcomings of the particular technique used in this instance to record degree of satisfaction with R.A.F. life. We will return to the first point later. To consider the second point, the form of the technique from which service satisfaction was assessed led subjects to make their judgements on an objective basis rather than a R.A.F. basis. They were invited to consider the satisfaction they found in their R.A.F. roles against a background of all roles of all men rather than only of men in the R.A.F. It is this that has probably caused the bunching at the 'satisfaction' end of the scale and so a poor scatter.

Two other lines of supporting evidence have to do with the rank composition of the two groups. It would seem reasonable to assume that officers and senior N.C.O.'s in the R.A.F. would tend to be more identified with the service than airmen. Officers and senior N.C.O.'s are older men who have decided to make a career of the R.A.F. and have had the reward of acceptance in a higher prestige group to confirm them in this decision. Airmen, on the other hand, are younger and are much less confirmed in a service career. Many of them will leave the service while still quite young men. Also they do not have positions of responsibility in the functioning of the service structure as the more senior men have, and are so less likely to feel closely identified with the service. It would seem safe therefore to assume that, on the whole, officers and senior N.C.O.'s are more likely to take their R.A.F. group as a positive reference than are airmen. Thus, it would be expected from our hypothesis that the positive mobility raters would tend to contain more officers and senior N.C.O.'s than the negative raters. And this is in fact found. Among the positive raters 68% were officers or senior N.C.O.'s, and among the negative raters 47%. This difference is statistically significant.

Table 1. *Relation of type of mobility rating to degree of satisfaction with R.A.F. life*

	Type of mobility rating		
Degree of satisfaction with R.A.F. life	Positive (%)		Negative (%)
Satisfied	71	—S—	52
Neutral	24		31
Dissatisfied	5	—S—	17
	100%		100%
N=	38		72

—S— = statistically significant difference $P = 0{\cdot}05$.
'Positive mobility rating' = exaggerated upward or underestimated downward mobility.
'Negative mobility rating' = exaggerated downward or underestimated upward mobility.

Another argument along the same lines has to do with the social status rating the two groups gave to the role of R.A.F. officer. It might be expected that men who feel closely identified with an institution would tend to rate the social status of the senior role in the institution more highly than would men less identified with it. The social status of such a role is in part a symbolic rating of the institution as a whole. In our two mobility groups 70% of the positive raters placed the role of R.A.F. officer in one of the highest two social status groups, while 50% of the negative raters did so. This difference is statistically significant and is in the direction to be expected from our hypothesis.

Finally, there is the analysis of the manner in which the two groups produced the abnormal ratings of their mobility. This is set out in Table 2.

There are several points of interest in this analysis. Consider first the ratings of the positive mobility raters. It can be seen that the abnormal ratings of these men arise mainly because they have rated their R.A.F. roles higher than the standard rating. These men have also underrated their native status significantly more frequently than have the neutral mobility raters. Among the negative mobility raters the opposite trends are present. They have abnormal ratings mainly because they have overrated their native status. They also underrated their R.A.F. roles significantly more frequently than the neutral mobility raters. All these trends are in line with our hypothesis.

VI. Summary

This paper set out to explain why two groups of men in the R.A.F. who otherwise had shown that they could assess social status differences fairly accurately did not assess accurately their own social mobility in moving from civilian life into the R.A.F. A hypothesis was set up which suggested that the men who exaggerated their upward mobility or underestimated their downward mobility did so because their R.A.F. role had a positive reference for them, while for the men who underestimated their upward mobility or overestimated their downward mobility their more important reference group was civilian. This assumed that when a man assesses the social status of a role which is symbolic (for him) of a positive reference group he would tend to raise its social status above its standard rating. And conversely, he would tend to depress the social status of a negative reference group. Several lines of evidence were considered which supported the suggestion that the men with positive mobility ratings were more identified with the R.A.F. than were the men with negative mobility ratings. Finally, an analysis of the rating behaviours of the two groups showed that the positive raters had a very much greater tendency to raise the status of their R.A.F. roles, while the negative raters showed an equally abnormal tendency to raise the status of their civilian backgrounds. Associated with this there was also a tendency to depress the complementary role, either civilian or service.

Table 2. *Analysis of abnormal status ratings of men making different types of mobility ratings*

	Type of mobility rating				
Deviant status ratings	Positive (%)		Neutral (%)		Negative (%)
Own R.A.F. status too high	60	—S—	23		14
Native status too low	29	—S—	4		3
Own R.A.F. status too low	3		6	—S—	30
Native Status too high	13		12	—S—	62
N =	38		84		72

—S— = statistically significant difference $P = 0{\cdot}05$.
All differences between corresponding percentages in the two outside columns are statistically significant.
A man may be counted in 2, 1, or 0 rows; thus column totals may be more or less than 100%.

Although the patterns of behaviour crucial to the testing of the hypothesis all showed significant trends in the expected directions, they were only trends and not complete conformities. From this it must be concluded that although the hypothesized tendencies have been operative, they do not furnish a complete explanation. It is submitted, however, that the material offers a sufficient empirical basis for a strong warning against the effects of reference group associations when dealing with ratings of social status.

REFERENCES

HALL, J. & JONES, D. (1950). Social grading of occupations. *Brit. J. Socio.* I, no. 1.
HYMAN, A. H. (1942). The psychology of social status. *Arch. Psychol., N.Y.*, no. 269.
KELLEY, H. H. (1952). Two functions of reference groups. *Readings in Social Psychology*, pp. 410–14. (Eds. G. E. Swanson, T. M. Newcomb, and E. L. Hartley.) New York: Henry Holt and Co.
NEWCOMB, T. M. (1943). *Personality and Social Change*. New York: Dryden Press.

(*Manuscript received* 18 *October* 1953)

ON THE VARIABILITY OF THE SPEED OF TALKING AND ON ITS RELATION TO THE LENGTH OF UTTERANCES IN CONVERSATIONS

By FRIEDA GOLDMAN-EISLER

Unit for Research in Occupational Adaptation, Medical Research Council, Maudsley Hospital

I. *Introduction* (p. 94). II. *The problem* (p. 94). III. *The material* (p. 95). IV. *The method* (p. 95). V. *The results* (pp. 95–104): (i) *Frequency distributions of speech rates* (pp. 95–96); (ii) *Individual differences and situational determination of speed of talking (speech rate)* (pp. 97–99); (iii) *The relation of speed and length of utterance* (pp. 99–101); (iv) *Length of utterance and variability in speed of talking* (pp. 101–103); (v) *Length of utterance and mean speed* (pp. 103–104). VI. *Discussion* (pp. 104–106). VII. *Summary* (pp. 106–107). *References* (p. 107).

I. Introduction

In a previous study of language behaviour (Goldman-Eisler 1954) twelve measurements of the temporal as well as of the grammatical aspects of conversation had been tested by the author for consistency in individuals. The type of conversation, which was a diagnostic interview, was kept constant. Among these measurements, a number of which turned out to be discriminating, the speed of talking (measured as the number of words a minute) was not only the most consistent, but also proved to be the most sensitive aspect of linguistic behaviour.

The measurements of speed of talking in the above experiment, however, were averages of the speed of word output of each individual during a complete interview, i.e. the total word output of the subject was divided by the total time spent talking. The interviews lasted fifteen minutes to half an hour, with an entire word output of 2000–3000 words by each speaker.

II. The problem

The present paper is concerned with the variability of speed of talking in the course of one speech situation for the different separate utterances. The problems are:

(1) How does the speed of talking of one individual vary within a single conversation for the many separate utterances which he makes in the course of it?

(2) Does speed of talking discriminate between individuals, and what is their variability from conversation to conversation?

(3) What is the relation of speed of utterance to length of utterance? By an utterance or speech action (these two terms will be used interchangeably) is meant a period of speech between the last word of the interlocutor's preceding utterance and the first word of his following one. In the present study speed is measured as the number of syllables a minute, as syllables are more appropriate units than words with which to measure the length of an utterance (Stetson, 1928).

III. THE MATERIAL

The material for answering the above questions was obtained from ten conversations and group discussions involving four normal adults (three men and one woman), who were all members of the academic and scientific staff of Maudsley Hospital, and four young female neurotics, interviewed each separately on a single occasion by the experimenter. The four normal subjects will be referred to as A, B, C, D and the four patients as L, M, N, O. Nineteen single records of A, B, C, D were derived from the following situations:

A talking with B, C, D separately, in a group of three (A, B, C), and in a group of four (A, B, C, D).
B talking with A, C, D separately, in a group of three (A, B, C), and in a group of four (A, B, C, D).
C talking with A, B, D separately, in a group of three (A, B, C), and in a group of four (A, B, C, D).
D talking with A, B, C separately, and in a group of four (A, B, C, D).

The material thus consists of twenty-three interaction records. Of these, twelve are records of paired conversations between normals, seven are records of group discussion among normals, while four are interview records between each of the patients and the experimenter.[1]

IV. THE METHOD

The length of each utterance was timed, while the speech recordings were transferred to paper and the number of syllables counted. The speed of talking for each utterance was then expressed as a rate of the number of syllables a minute. The total number of utterances recorded for A, B, C, D (i.e. the normals) was about 1299, ranging between 300 and 400 for each individual, with a mean number for each conversation ranging between 50 and 100 utterances. The length of these utterances varied from one to many hundreds of syllables. For L, M, N, O (i.e. the neurotics) there were altogether 376 utterances, with the number of utterances per interview ranging from 78 to 104. Every conversation and group discussion lasted on an average about three-quarters of an hour, and the interviews between half an hour and three-quarters of an hour each.

V. THE RESULTS

(i) *Frequency distributions of speech rates*

Frequency distributions were plotted for the talking speeds of each individual separately as shown in Fig. 1. All of these distributions are more or less positively skewed, and are also more or less leptokurtic. There are individual differences in the degree of leptokurticity (peakedness) and in the skewness of the curves representing the talking speeds derived from the individual conversations, but the general description holds for all distributions in the sample in respect of skewness and for six out of the eight distributions in respect of leptokurticity. As an indication of skewness it may be mentioned that the numbers of positive readings beyond $+2$ S.D. in relation to the negative readings below -2 S.D. are in every case but one a multiple of the latter, the proportion of positive readings surpassing the $2\frac{1}{2}\%$ of total frequencies which one might expect in a normal distribution while the number of negative readings below −2 s.d. is considerably less than 2½% of the total frequency. On the average 1% of the total frequency occurred below −2 s.d. and 4·6% above +2 s.d. The peakedness of the distribution is described by comparing the interquartile ranges of the distributions within which the middle

---

[1] Twelve more records of patients in diagnostic interviews with a psychiatrist have since been analysed, confirming our conclusions.

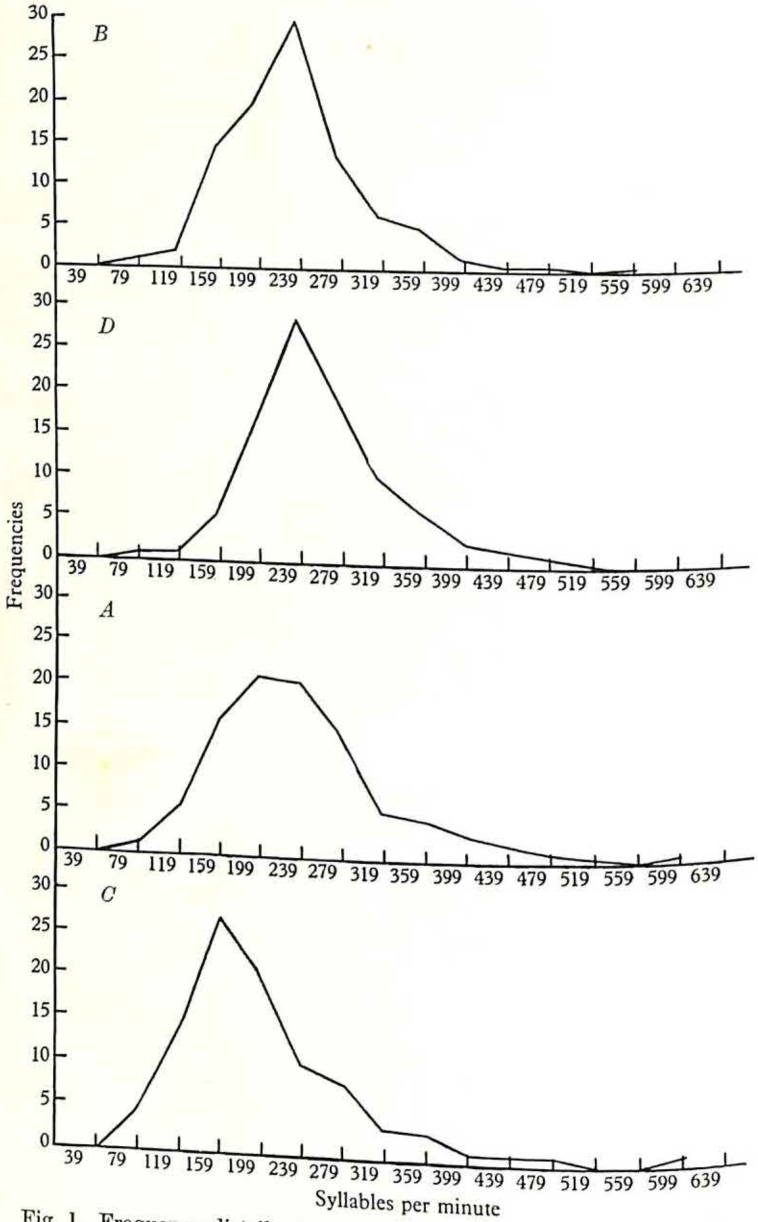


Fig. 1. Frequency distributions of speech rates of subjects *A*, *B*, *C*, *D*.

50% of the frequencies lie with the ranges lying between ±0·674 s.d. and ±1 P.E. which would contain 50% of the frequencies if the distribution curves were normal. In six out of the eight sample distributions the middle 50% of the frequencies were concentrated within narrower limits than expected; these curves were clearly leptokurtic as well as skewed. In the case of *C* and *O* the middle 50% lay near enough between the normal limits of ±0·674 s.d. These curves were skewed but normal in respect of peakedness.

(ii) *Individual differences and situational determination of speed of talking* (*speech rate*)

The subjects *A*, *B*, *C*, *D* and *L*, *M*, *N*, *O* were compared in respect of their speed of talking. Comparisons were made separately for the different lengths of utterance. For *A*, *B*, *C*, *D*, measurements from all their conversational situations were pooled, and in the case of *L*, *M*, *N*, *O*, the measurements were taken from their one interview. Analyses of variance were then computed.[1]

Table 3 illustrates clearly, in terms of means as well as standard deviations, how the speed of talking (speech rate) gradually gains stability as utterances grow longer. Speech rate discriminated between *A*, *B*, *C*, *D* at a high level of significance ($P=0{\cdot}001$) over the whole range of the lengths of utterances from very short to very long ones.

The records of the patients, despite the fact that there were fewer measurements available, showed this differentiation for all but the shortest length of utterance. It is also remarkable that those individuals who were fast talkers in their short utterances remained so in the long ones, notwithstanding the general decrease in speech rate with increase of length of utterance. The rank order of the subjects in respect of speed of talking remained consistently the same over the whole range of lengths of utterance.

Table 1. *Means and standard deviations of speech rates* (*number of syllables per minute*) *for subjects A, B, C, D* (*normals*) *and L, M, N, O* (*patients*)

	Mean*	No. of syllables	S.D.	*P*		Mean	No. of syllables	S.D.	*P*
D	282·7	374	84·3	0·001	*M*	253·5	77	80·0	0·001
B	257·1	274	77·0		*O*	245·9	104	68·6	
A	247·3	295	86·8		*L*	222·1	104	70·5	
C	208·3	357	56·2		*N*	214·8	78	88·9	

* The means for *A*, *B*, *C*, *D* are based on readings from all their conversations.

After excluding the short utterances of less than 40 syllables length on account of their wide scatter, measurements of talking speed (speech rate) taken from utterances over 40 syllables long were submitted to an analysis of variance. The mean speech rates in each of the conversations were also compared for each of the four subjects *A*, *B*, *C*, *D* separately. Table 2 lists their means and standard deviations. From it the reader may see that the speed of talking, besides characterizing individuals over and above their fluctuations on different occasions (Table 1) may differ significantly for any one individual when compared for each situation separately.

The nature of the data, being unequal in numbers as well as errors for the groups to be compared, did not favour a three-way analysis of variance sorting out the relative contributions of individuals' speech situations and their interaction, simultaneously. But it was obvious from the separate analyses of variance that all three were significant sources of variation for the speed level of talking in conversations, in proportions varying with the different occasions and individuals. The degrees of variability of individuals from occasion to occasion seems, however, to vary with the individual. Thus *B*'s range of variation is much more flexible than that of *A*, *D*, and particularly *C*. The table also gives an indication of interaction between individual speakers and particular conversations. No systematic effects of the individual interlocutors for all their partners can be observed, although the discussion in which all four subjects took part accelerated all of them.

---

[1] It is well known that a slight departure from symmetry does not affect the test of significance performed by an analysis of variance very much (Cochran & Cox, 1950).

The complexity of factors determining the speed of talking is illustrated in the following instance. *C*, while having an accelerating effect on *A* in one conversation, did not have this effect in the other, nor when she joined the conversation that had taken place between *A* and *B* alone, as a third in the group. When talking to *B* alone, *A* spoke 216·7 syllables a minute on the average (the total duration was about 30 min.). *C* then joined the group and they continued as a group of three. *A* spoke if anything at a slower rate (209·5) with *C* present. *B*, on the other hand, who had reached his lowest rate in conversation with *C* alone, spoke significantly faster after *C* had joined the group.

Table 2. *Results of analysis of variance testing the differences in speech rates in different conversational situations for subjects A, B, C, D*

Subject	Conversation	Means	S.D.	*F* ratio	*P*
A	With *B*	216·7	57·5	4·8	0·01
	With *C*	254·5	58·3		
	With *D*	206·0	80·4		
	With *B*, *C*	209·5	23·2		
	With *B*, *C*, *D*	244·8	24·4		
B	With *A*	206·1	31·3	16·3	0·001
	With *C*	193·8	36·5		
	With *D*	252·2	56·9		
	With *A*, *C*	237·2	52·5		
	With *A*, *C*, *D*	275·7	42·9		
C	With *A*	181·0	46·1	1·5	N.S.
	With *B*	186·6	29·1		
	With *D*	187·5	48·5		
	With *A*, *B*	206·2	96·4		
	With *A*, *B*, *D*	214·5	39·6		
D	With *A*	244·8	52·1	3·18	0·05
	With *B*	273·0	44·3		
	With *C*	272·7	87·3		
	With *A*, *B*, *C*	274·1	41·3		

Conclusion

The only general conclusions suggested by our data so far are that within any one conversation individuals tend to maintain a certain level of speech rate, which may be an indication of a general level of stimulation either aroused on the occasion of this conversation, or existing at the time and maintained during its course. This level of speech rate is specific to the occasion, varying for any one individual from single occasion to single occasion, whereby the person of the interlocutor or the combination of a group are only one factor in the variation. (That it might at times be an important factor is suggested by the fact that all speakers accelerated when taking part in a general group discussion with *D* present.)

Other factors have so far not been investigated, but one would imagine that the level of speech rate on any one occasion is a reflexion of a number of stimuli, apart from the interlocutor, and that the effect of the latter would have to be broken down into effects of his social status in relation to the speaker, of his personality in general, and of his speech behaviour in terms of what he says as well as of formal characteristics of speech. The stimuli, apart from the interlocutor might include biological wellbeing, mood, interest aroused by the subject of conversation, as well as the degree of his familiarity with it.

However, over and above its fluctuations from occasion to occasion, speech rate differentiates between individuals. All the subjects investigated have shown themselves relatively limited in the degree to which they raised or lowered their speed of talking for any length of time. The fast speakers remained consistently fast and the slow speakers consistently slow.[1] This is in keeping with the author's previous results based on doctors and patients in psychiatric interviews.

(iii) *The relation of speed and length of utterance*

The third question asked was what bearing the length of an utterance (or the number of syllables spoken) has on the speed of talking. Having excluded the very short utterances, i.e. those consisting of less than five syllables (it is obvious that rates of speech output for these could not hope to be accurate), the speeds of the separate utterances during each conversation were plotted on the abscissa against their respective lengths. The scale for the latter on the ordinate is in terms of the number of syllables contained in each utterance.

There were twenty-three scatter diagrams in all, containing an average of sixty-eight measurements (ranging from 49 to 93) per function. The type of relation which emerged is illustrated in Figs. 2–5, each being a combined scatter diagram for every individual using the readings from their several conversations.

The characteristic feature of these functions, namely their shape, is common to all twenty-three diagrams underlying them. The variations lie (*a*) in the location of central values of speeds on the abscissa, (*b*) in the width of scatter, and (*c*) in the height of the column, i.e. in the length of utterances to which each individual rises. The common elements are the following:

(i) Speeds for short utterances up to 20 and 40 syllables contribute most to the base of the curve. The latter spreads over the range from about 60 to 600 syllables a minute for the shortest utterances, i.e. from very slow to very fast, and narrows down as they grow longer.

(ii) Stability is gained at a length of utterance of about 100 syllables. At a length of utterance of 100 syllables the area of the curve has finally narrowed with all the utterances of greater length falling within a column.

(iii) Speed in short speech periods, while ranging from the extreme left to the extreme right of the distribution, is responsible for its skewness to the right. The right tail, as well as containing more measurements, is far longer.

(iv) High speeds of talking rarely occur in the case of utterances more than 60 syllables in length, and *never* in the case of those exceeding 100 syllables. The curve as a whole recedes towards the central value when the length of utterance is between 40 and 60 syllables.

Table 3 gives the means and standard deviations for different lengths of utterance (i.e. utterances containing 5–19, 20–39, 40–59, 60–79, 80–99, 100–199, and more than 200 syllables (this last includes utterances of up to 813 syllables in length). Figs. 5 and 6 show the diminution in the rapidity of speech and in the range of speeds for utterances of increasing length. In spite of the wide scatter along the scale of speed in the short utterances, speech rates are not evenly spread there, a central tendency being clearly discernible. An analysis of variance computed for the purpose of finding out at which length of utterance speech rate could discriminate between individuals (taking measurements from their entire conversation) showed that even in the very short utterances there

---

[1] Reading speeds based on five English tests of various types (prose and poetry) taken from *A*, *B*, *C*, *D* show this rank order of speech rate to be the same as for talking in conversations.

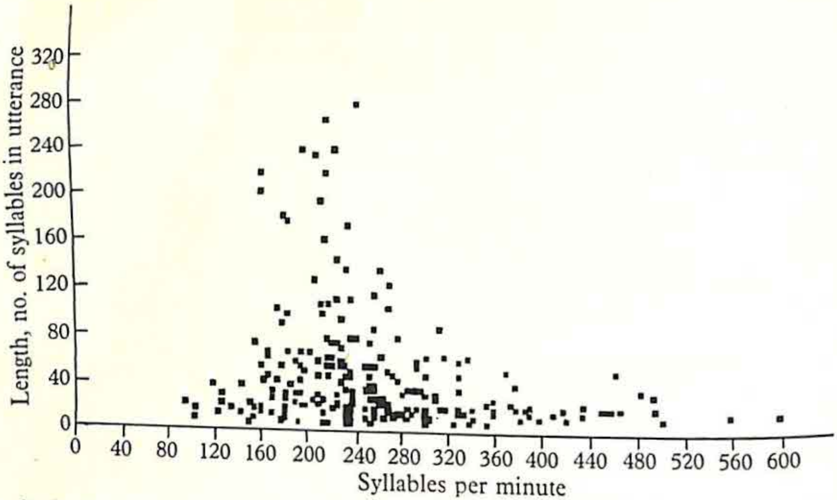


Fig. 2. Scatter diagram showing relation of speed and length of utterances for subject *A*.

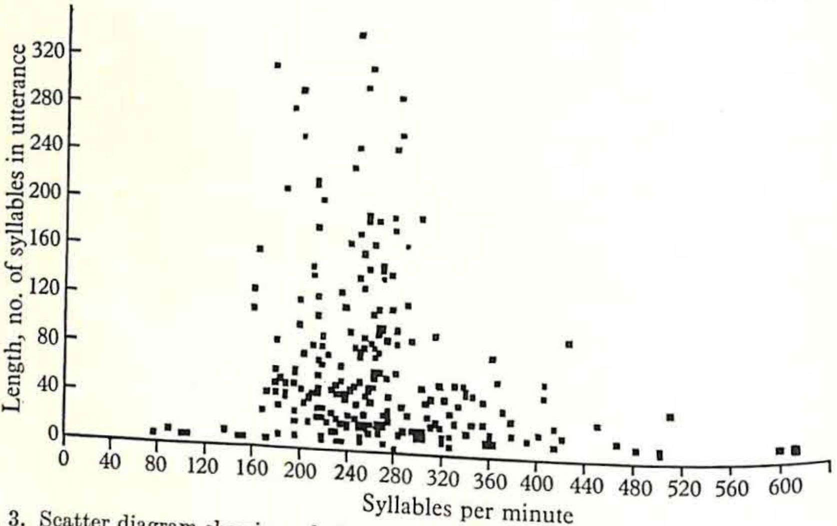


Fig. 3. Scatter diagram showing relation of speed and length of utterances for subject *B*.

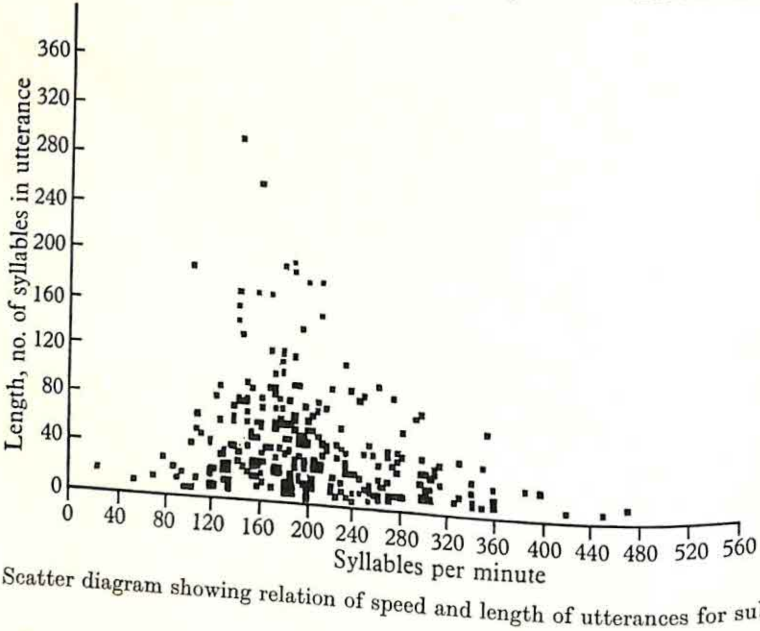


Fig. 4. Scatter diagram showing relation of speed and length of utterances for subject *C*.

was enough concentration around the central tendency for mean rate of speech to discriminate between the individuals *A*, *B*, *C*, *D* at a highly significant level of probability (*P* 0·001).

Testing the discriminative capacity of speech rate for utterances of increasing length for the four records of the patients' interviews (this measurement came from only one

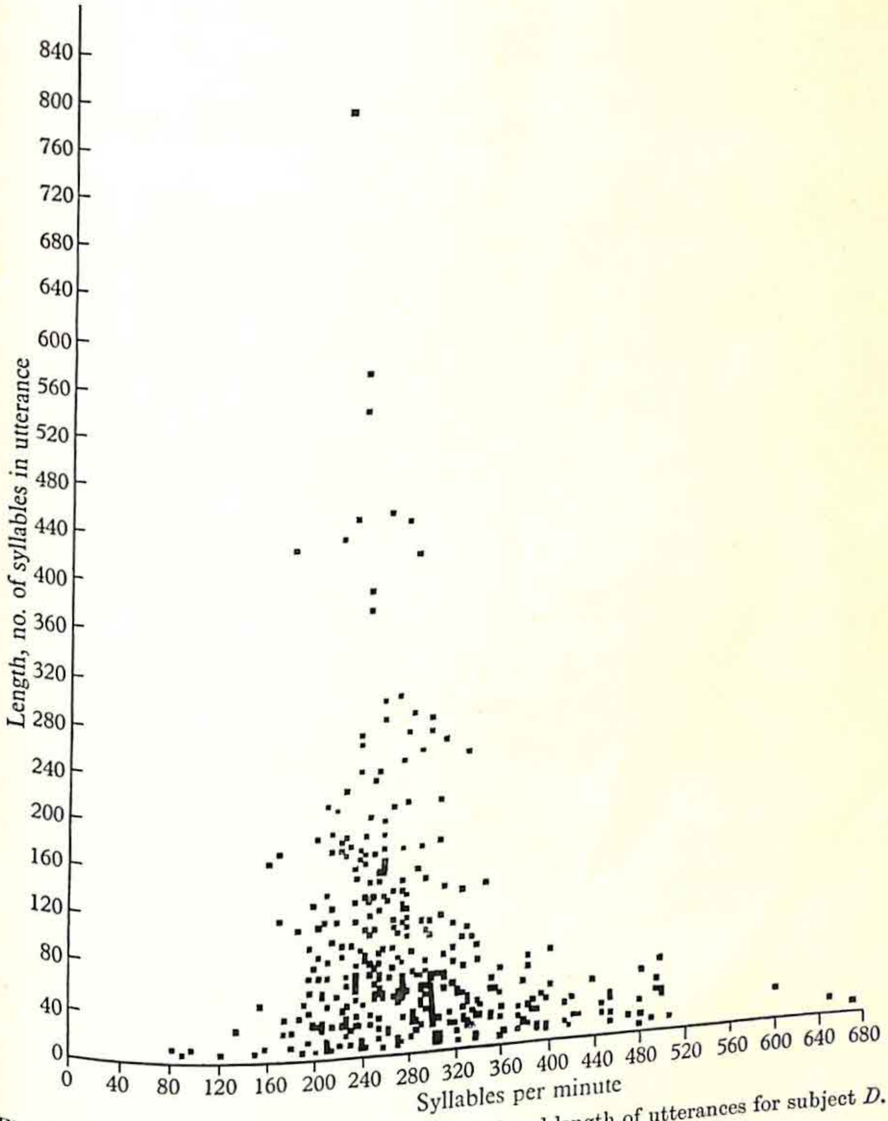


Fig. 5. Scatter diagram showing relation of speed and length of utterances for subject *D*.

interview), the mean speech rates for short utterances not exceeding 20 syllables proved incapable of discriminating between the four patients. But the speeds of talking for utterances of between 20 and 40 syllables in length, and also for those still longer, discriminated significantly as was mentioned above (Table 1).

(iv) *Length of utterance and variability in speed of talking*

Fig. 6 shows the pooled standard deviations for utterances of different length of the members of the hospital staff *A*, *B*, *C*, *D*, and separately for those of the four patients

L, *M*, *N*, *O*. The regression line drawn through the points of the two curves shows a virtually identical slope, the standard deviation decreasing as the number of syllables uttered increases up to 200, at a rate of about 7 : 12. From there on the standard deviation

Table 3. *Means and standard deviations of speech rates (number of syllables per minute) for different lengths of utterances*

	Syllables in utterance...	5–19	20–39	40–59	60–79	80–99	100–199	Over 200
A, *B*, *C*, *D*	General mean	272·3	254·2	239·2	233·2	237·6	230·8	229·8
	S.D.	108	79·7	71·2	56·4	53·9	40·8	31·0
	No. of utterances	319	315	218	119	84	165	78
L, *M*, *N*, *O*	Mean	264·6	244·4	219·1	207·9		206·5	
	S.D.	99·7	53·9	45·6	31·3		30·0	
	No. of utterances	109	85	46	50		73	

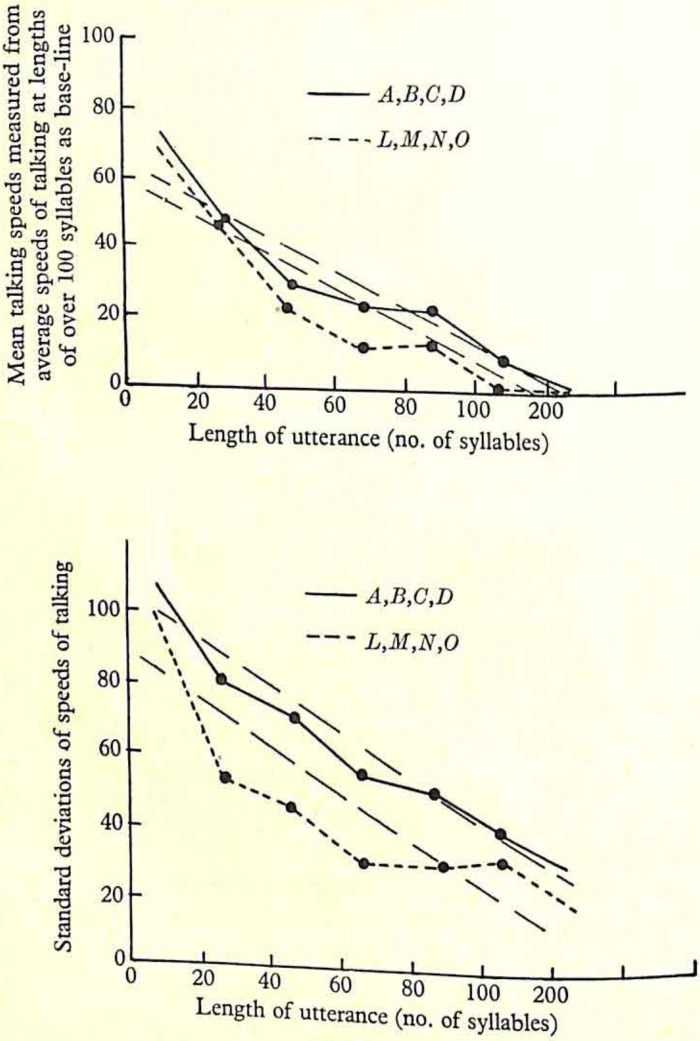


Fig. 6.

remains relatively constant. This is so in the case of the normals, *A*, *B*, *C*, *D*, as well as in that of the patients, with the difference that for the latter the scatter diminishes more rapidly, and constancy is attained with shorter utterances than in the case of the normal group. Whether this is due to the fact that a greater number of measurements exists for the normals, or whether it may be a characteristic of neurotics that they are unable to maintain high speeds as long as normals is for the present a matter of speculation. However, the unequivocal nature of the general tendency of this relation should enable whoever wishes to investigate the question to answer it with a high degree of precision.

(v) *Length of utterance and mean speed*

As mentioned above, a similar decrease with increasing length of utterance was observed in the means of speech rates calculated separately for utterances of the following lengths: ontaining 5–19, 20–39, 40–59, 60–79, 80–99, 100–199, and more than 200 syllables. Again, this was so for all records of normals and patients alike, irrespective of the individual differences in mean speech rates which had proved highly significant. When mean speech rates were plotted for utterances of different length the slope of the regression line again showed little variation for the different individuals. To demonstrate this, measurements of mean speech rates were again pooled into two groups, one for normals and the other for patients. The different speed levels for the different individuals were equalized by taking the mean talking speed for the longest utterances (over 200 syllables) for each individual as his base-line or zero-point, which was subtracted from mean talking speeds of all the shorter utterances (six more classes of utterances). Fig. 5 shows that not only the angle of the slope but the shape of the curve of means as they decrease towards the basic mean speech rate (i.e. the talking speed with the greatest stability or smallest standard deviation) is remarkably similar in its characteristic way for normals and patients. It should, of course, be kept in mind that the high mean speeds of the short utterances cover a very wide range and that only for the longer utterances are the mean speech rates representative of the real values. As to the shape of the curve it seems as if the utterances of different length fell into three or four classes: (1) the very short utterances of up to 20 syllables and the next short ones of up to 40 syllables length with mean speech rates of 272·3 syllables a minute, and 254·2 syllables a minute, with a scatter of S.D. = 108·0 and 79·7; (2) utterances of middle length, that is those containing between 40 and 100 syllables, whose mean talking speeds seem to form a plateau in the curve and whose scatter has considerably decreased in the same way; and (3) long utterances of over 100 syllables' length, with a somewhat slower rate of talking having gained final stability. These groupings suggest that the different types of utterances may represent different functions in interaction. For the most part, retorts, expostulations and interjections would seem

Table 4. *Means and standard deviations of rates of talking for different lengths of utterances based on all speech samples recorded of A, B, C, D, and on interview recorded of L, M, N, O*

	A, B, C, D			*L, M, N, O*		
Length of utterance	No. of utterances	Mean	S.D.	No. of utterances	Mean	S.D.
5–19	319	272·3	108·0	109	264·6	99·7
20–39	315	254·2	79·7	85	244·4	53·9
40–59	218	239·2	71·2	46	219·1	45·6
60–79	119	233·2	56·4	50	207·9	31·3
80–99	84	237·6	53·9			
100–199*	165	230·8	40·8	73	206·5	30·0
200 and over	78	229·8	31·0			

* Owing to the decrease in the frequency of utterances with increase in length, the speech rates underlying the means of utterances longer than 100 syllables were grouped into larger intervals.

to make up the two groups of short utterances, conversational exchanges the middle group of utterances, and the more considered statements, reports, deliberations, outpourings or diatribes of greater length would constitute the last group. It should be mentioned in this connexion that individuals differed considerably in the proportions of their total number of utterances assignable to the groups of utterances listed above It so happened that in our sample the normal as well as the neurotic subjects divided, in respect of length of utterance, into two fairly distinct types of talkers, as Table 5 shows. *A* and *C*, whose utterances are mainly short, are also the slow speakers in the group; whereas *B* and *D*, who usually speak at length, are also those with considerably higher mean speeds of talking (*D* 265·6 and *B* 240·0, *A* 226·5 and *C* 191·3). It is also of interest that the differences lie not in the utterances of medium length but in the short and long ones, *A* and *C* concentrating on the former, i.e. mainly responding and interjecting, while *B* and *D* show a much greater degree of staying power in talking.

The corresponding distribution of utterances of different length for the four patients showed a similar dichotomy, as shown in Table 5.

Table 5. *Proportions of utterances of different lengths in total output of A, B, C, D and L, M, N, O (in percentages)*

Length of utterance	*A*		*B*		*C*		*D*	
5–19	33	57	20	40	29	58	18	40
20–39	24		20		29		22	
40–99	33	43	32	60	33	42	25	60
100 and over	10		28		9		35	
	L		*M*		*N*		*O*	
5–19	20	46	23	41	37	56	38	63
20–39	26		18		19		25	
40–99	29	54	26	59	24	43	23	37
100 and over	25		33		19		14	

VI. DISCUSSION

To gain a better understanding of the implications of the phenomena described above it may be useful to contemplate the uniform way in which the rate of talking fell off markedly as utterances became longer. The fact that on no occasion were longer utterances spoken at high speeds suggests that a physical as well as a psychological factor is here operating as an impediment; in other words, the high rates of word output sometimes achieved in short bursts seem, under ordinary circumstances, physically impossible to maintain for any length of time. As we have seen, short utterances were not necessarily always faster than long ones, but in the case of the former the range of speed was wider from very slow (60 per min.) to very fast (600 per min.). What discriminates the short from the long utterances then is not only the rates at which they are spoken, but the inherent difference in the amount of freedom to vary speed. The latter is partly dependent on the physiological mechanism involved in speech production. Respiration in particular, which is the source of energy maintaining the output of speech, must be taken into account (Stetson, 1928). The production of speech requires a renewed column of exhaled air for extended vocalization and articulation. While the rate and rhythm of respiratory movements can be modified at will, the exercise of voluntary control over breathing is strictly limited by the respiratory reflexes which serve the more vital and primary functions of the organism (Travis, 1931). As it is biological necessity to interrupt the stream of speech at more or less regular intervals in order to get air into the lungs, repeated inspirations which, however abbreviated, take time, will affect the durations of long utterances.

Utterances of short duration, on the other hand, may to a considerable degree escape the retarding effect of repeated respiratory action, and gain time by suspending breathing for short periods.

Apart from this physiological limitation to the pace at which speech is produced, there are mental processes involved in the production of speech; namely the need for conceptualizing, integrating, and phrasing, in other words the process of thinking involved in utterances of greater length which is generally more complicated. While the economy of breath control on the physical side, and the integrative processes of language formation on the mental, are factors setting limits to the rate of word output, factors having the opposite effect are also suggested. Examination of the records of the talking speeds of partners in conversation showed that utterances spoken at high speed by one speaker either followed or were followed by fast utterances, or started a series of quick exchanges. Another occasion during which the rate of talking was speeded up greatly was when an interlocutor interjected a remark into a breathing or thinking pause in the stream of an utterance before the speaker had come to an end. In such a case, which was a frequent occurrence in the discussions of *A*, *B*, *C*, *D*, the usual consequence was that after the interruption the rate of talking rose, often quite steeply, only to slow down again as the interlocutor relapsed into passivity. The speaker who interrupted, unless his interjection was very short, also, as might be expected, spoke faster than usual, having to fit his utterance into the pause of the other's stream of words.

The difference in the conditions of interaction under which short interjections or lengthy statements are made may be described as one between participation and sole possession of the field. And the object of participation or possession in speech situations is time, the dimension in which linguistic behaviour progresses. In a conversational situation, time is a quantity to be shared between the partners, and the more spontaneous such a situation is, the more open is left the manner (i.e. the more subject to the interaction of the interlocutors) in which this time for talking will be distributed between them. The problem of the distribution of time in interactional situations becomes critical when the demand for it exceeds the supply, i.e. when the partners in a conversation have the urge to speak, either simultaneously or when the speaker's urge for talking is so disproportionate as to prompt him to encroach beyond the quantity of time which his partner is prepared to forego. The fact that in our sample those who said most also spoke the fastest, seems from this angle to become easily understandable. For loquacious persons, time is in short supply, particularly when their conversational partners also have a lot to say. Only by increasing their word output per time unit can they manage to release material pressing for verbal expression. The acceleration in an individual's rate of talking may thus be taken as a measure of the intensity of this pressure towards verbal expression. Where both partners in a conversation seek this release simultaneously, the result is a clash of conversational actions, i.e. simultaneous talk, rapid bursts of speech, a series of interruptions and quick short exchanges. The clash manifesting itself in these forms is over the possession of time, or more precisely of a particular stretch of time. The circumstances in which acceleration in the rate of talking was observed, suggests that it is a particular form of action by which either the speaker's time is defended against his partner's encroachments, or the co-speaker's time appropriated or invaded while he still wishes to continue talking. That an increase in the number of interjections and a cut in the time available can have the effect of speeding up the rate of talking to a spectacular degree has already been demonstrated in other research (Goldman-Eisler 1954). At which point time appears to a speaker to be in short supply is also a matter for his own subjective assessment, and the speed of talking would be an indication of subjective time as experienced by the speaker. This, one might assume, will be a function of the quantity of material that is urging to be expressed, the degree of urgency felt, the objective amount of time available, and, more basically, of time perception as determined by physiological and psychological processes (Lecompte de Nouy, 1936; Schilder, 1936).

In the light of the above considerations it would appear, then, that short utterances in conversation being relatively unhampered by the functional necessities of the physical and mental processes involved in speech production, and therefore being free to cover a very wide range of speeds, will have a much greater capacity than long utterances to reflect the intensity of interaction.

Long utterances, on the other hand, represent a different aspect of the conversational situation. They are the result of a speaker having taken the floor, and are dependent on sufficient time being available. They are conditional, too, on the passivity, whether voluntary or enforced, of the interlocutor, enabling a speaker to continue talking at some length. The speaker thus talks, so to speak, in his own time.

It is the speed of talking in the long utterances, therefore, which indicates to what extent a speaker considers time to be his 'own'. In other words, the speed of talking in long utterances which has emerged as having great stability within any particular conversational situation, and as being characteristic for individuals, although flexible with varying situations, may turn out to be a measure of an individual's subjective time in any particular situation, i.e. of how much time he feels he has, and how secure he feels in its possession.

Further investigation is, of course, necessary to corroborate the suggestion of this link between rate of talking and time sense.

VII. SUMMARY

1. The speed of talking (or speech rate) had previously been shown to be a consistent as well as a highly sensitive aspect of conversational behaviour.

2. This paper deals (*a*) with the question of the speed of talking measured as the number of syllables a minute for the separate utterances as distributed within a single conversation, (*b*) with individual differences and the situational determination of the rate of speech, and (*c*) with the relation of speed of utterance to length of utterance.

3. The material was based on transcribed speech recordings of ten conversations and group discussions involving four normal adults (three men and one woman) of academic status and four interviews of young female neurotics yielding altogether twenty-three records. The conclusions are based on a total number of 1299 utterances.

4. The frequency distributions of speech rates for the different individuals in different situations had the common characteristic of being skewed positively, and six of the eight distribution curves were leptokurtic.

5. Speech rate discriminated between individuals at a high level of significance. It fluctuated for any one individual at varying degrees in the different situations, but within the intra-individual range. No systematic effect on speech rate was exercised by

persons, the interlocutor emerging rather as *one* of the factors affecting the speed of talking.

6. The relation between speed and length of utterance was illustrated by scatter diagrams for each subject, based on pooling the utterances from all their conversations. They show that the speeds of utterances, which lie along the base of the distribution, are constituted entirely of short utterances, rarely containing more than 40 syllables, while the body of the distribution includes all the long utterances. Length thus impinges upon speed of utterance in the following ways.

(*a*) The range of speech rates is relatively great for short utterances, but narrows as the utterances grow longer.

(*b*) Stability is gained at a length of about 100 syllables. Speech rates for utterances longer than that are confined within the area of a column, its base being to the left of the mid-point, and its height representing lengths of utterances.

(*c*) Fast speech rates occur rarely at lengths above 60 syllables per utterance and *never* above 100 syllables.

(*d*) Means and standard deviations of speech rates for utterances of different lengths decreased with the increasing length of utterance: the stabilization of speed in talking being in the direction of a slowing down.

7. In spite of the wide scatter, speech rates for short utterances as well as long ones discriminated between individuals with high significance.

8. The curve of decreasing mean speeds for increasing mean lengths is in shape and slope very similar for the four normal and the four neurotic subjects. It points to three and possibly four types of utterances according to their lengths, suggesting as many different functions in interaction.

9. Further analysis as to the meaning of the rate of talking from its relation to length of utterance, pointed to a link between speech rate and sense of time.

References

Cochran, W. G. & Cox, M. G. (1950). *Experimental Design*, pp. 83–4. New York: John Wiley and Son.

Goldman-Eisler, F. (1954). A study of individual differences in the behaviour of some aspects of language in interviews. *J. mental Science*, no. 100, 177–197.

Lecompte du Nouy (1936). *Biological Time*. London: Methuen and Co.

Schilder, Paul (1936). Psychopathology of time. *J. Nerv. Ment. Dis.* lxxxiii, 530–46.

Stetson, R. H. (1928). Motor phonetics. A study of speech movements in action. *Arch. néerl. Phon. exp.* iii, 1–216.

Travis, L. E. (1931). *Speech Pathology. A Dynamic Neurological Treatment of Normal Speech and Speech Deviations*. New York: D. Appleton and Co.

(*Manuscript received* 9 *July* 1953)

AN EXPERIMENTAL STUDY OF COMPARATIVE JUDGEMENTS OF TIME

By JOHN COHEN, C. E. M. HANSEL AND JOHN SYLVESTER
University of Manchester

I. *Introduction* (p. 108). II. *Aim of the experiment* (p. 108). III. *Subjects and procedure* (p. 109). IV. *Results* (pp. 109–110). V. *Interpretation of results* (pp. 110–113). VI. *Consideration of hypotheses* (p. 113). VII. *Summary* (p. 114). *References* (p. 114).

I. Introduction

The seeming duration of outer or private events is perhaps the most fundamental dimension of human experience. In spite of this the literature on the subject is disappointing. Woodrow (1951) remarks in his recent review that data accumulated by various investigators have two outstanding characteristics, one being their mutual inconsistency and the other their 'mentalistic' nature. It is not clear whether he considers the second characteristic a merit or defect. The studies considered by Woodrow have been very largely confined to the comparison and reproduction of intervals up to about 10 sec. in duration, but there have also been introspective accounts of single intervals of three or four minutes in duration. Woodrow's review scarcely does justice to the subject, for he overlooks the basic problems of the psychology of time and the frequently reported temporal phenomena which are of universal interest, for example, the phenomenon that the years seem to pass more rapidly as we grow older (James, 1901). Nor does he refer to the only serious book (Sturt, 1925) entirely devoted to the psychology of time in which many of these questions are raised and discussed in the light of evidence of a kind which he does not consider.

A different approach to the study of time was introduced by Piéron (1952) and his students in France in 1929. Piéron's view is that units of subjective time are functions of the velocity of deep-seated organic processes which determine metabolism and other vital activities. As is well known, this approach has been developed by Hoagland (1935) in the United States, but important as the work undoubtedly is, it has only a limited bearing on the experience of time, principally in elucidating what might be called the organism's 'chronostatic' mechanisms. In the present study we attempt to explore the subjective duration of long intervals of time.

II. Aim of the experiment

Our aim in this experiment has been to discover the manner in which the seeming duration of time since different past events varies with the lapse of chronological time since these events. There is at least one hypothesis which might lead one to expect a non-linear relationship between chronological intervals, as they extend further and further into the past, and our judgements of them, such judgements presumably showing a relative contraction. On the basis of Weber's law we might expect quantitative judgements of our experience of time intervals, at any particular age, to increase linearly as the chronological intervals increase logarithmically.

III. SUBJECTS AND PROCEDURE

There were seventy-two children, thirty undergraduates, and 130 students of Adult Classes. These were divided into groups as shown in Table 1. All groups included both sexes. Twenty-two of the twenty-four children of mean age 13 years 6 months were studied under two sets of conditions and therefore appear twice in Table 1.

The essence of the procedure was that subjects marked off lengths of a line to represent duration of past intervals. Subjects experienced no difficulty in this procedure. This is not surprising if we consider that in everyday language words denoting length are used to refer to periods of time (e.g. a short time, a long time) and also that, in respect of short intervals of time, judgements of spatial and temporal intervals are inseparable (Cohen, Hansel & Sylvester, 1953).

The subject was given a line 10 in. long on a sheet of paper. He was asked to imagine that this line represented his life from 'birth' to 'now', and to mark off a length representing a given interval. A different line was given for judging each interval. Seven different intervals were presented in random order, this procedure being repeated four times. In the second and fourth presentations, the end-points of the line representing 'birth' and 'now' respectively were reversed. The subjects were carefully instructed to indicate, by marking off a length of the line, how long a given interval *seemed* in relation to the total past life or, as in part of the experiment, in relation to another fixed interval (e.g. 1 year).

In order to make our procedure clear, we must make a distinction between (i) an interval since an event ('How long does it seem since event *X* occurred?'), and (ii) the point on a time scale when an event occurred ('When did the event *X* occur?'). We are concerned with (i) only. In this case the subject is asked to indicate or represent his 'feeling' of the duration of an interval without conscious calculation of chronological time. It must be clear to him *when* the delimiting event occurred. In the second case, with which we are not here concerned, the subject would be asked to locate an event on a line representing his life, or part of it, to date, without necessarily knowing precisely *when* the event occurred.

The choice of birth as an anchoring point was made for two reasons, first it is definite and easily ascertainable, and secondly we assume that the whole of a person's past life is in an important sense a unity.

In the case of groups (v) and (vii) (consisting of students of Adult Classes) in Table 1, all the subjects were assembled together, and twelve intervals were presented in random order in a single series.

Because the experiments with the various groups had to be carried out on different occasions, and also because some intervals suitable for children (e.g. the interval since the end of the school holidays) were unsuitable for adults, the intervals chosen also had to vary.

The details relating to the groups and procedures are set out in Table 1.

IV. RESULTS

The means and standard errors are set out in Table 2, and the means are plotted against the logarithms of the intervals in Fig. 1. The estimates are lengths marked off a line representing the entire life. We have calculated the medians, in addition to the means because some of the distributions depart from normality and in such cases it may be argued that the means are not the most representative values of the variable. These medians are shown in Fig. 2. We return to this later.

V. INTERPRETATION OF RESULTS

The figures in Table 2 indicate comparative judgements of time intervals at each age. The distributions of the estimates for any interval given in the table vary markedly as between the children and the adults. The estimates of the children are approximately normally distributed, whereas those of the adults are skewed and in some cases bimodal. We may infer from this that the skewness and bimodality are not attributable to the scale, and that under favourable conditions the estimates yield a normal distribution. The skewness and bimodality of the distributions for the adults may be due to the fact that in the older groups, a proportion of subjects make their estimates on a linear basis, as a result perhaps of their experience of the use of linear scales. As stated above, it is because of this skewness that the medians have been calculated. It should be noted, however, that the medians may be unduly influenced by the proportion of subjects making linear estimates.

Table 1

Group	Number	Age		Method		Scale
		Yr.	Mnth.	No. of intervals	Times presented	
(i)	25	8	3	7	2	10 in. = life
(ii)	23	11	10	7	2	10 in. = life
(iii)*	22	13	6	7	4	10 in. = life
(iv)	30	19		7	4	10 in. = life
(v)†	130	20–74		11	1	10 in. = life
(vi)	24	13	6	7	4	10 in. = 1 year
(vii)‡	7	46		8	1	12 in. = 10 years

* This group is also included in group (vi).
† In this group one interval was repeated as a check.
‡ The estimates are plotted in Fig. 4.

For intervals up to about 6 months the form taken by the subject's estimates is such that as the actual intervals increase logarithmically his estimates increase in approximately linear fashion. Intervals beyond 1 year yield estimates which are approximately linear and go back to the point representing birth. In other words, the linear portion of the line is divided up in a way which would be arrived at by calculation. This linear effect is, indeed, obtained when subjects instead of being asked to judge apparent intervals, are asked to locate historical events along the given line. It would seem, therefore, that by this method the subject is unable to appreciate comparative intervals greater than about 1 year from now and is only able to locate events.

Thus, the immediately preceding 24 hr., for example, seems a relatively long time in proportion to the preceding month, whereas there is no perceptible disproportion when comparing the previous 5 years with the previous 10 years. These latter judgements (i.e. for intervals beyond 1 year) are made in an entirely different fashion and seem unable to be formed without reference to dates or events in one's personal past to which dates can be assigned or age estimated.

Estimates of intervals between 6 months and 1 year are highly variable. It would seem that the estimates are sometimes arrived at by the method of shorter intervals and sometimes by the method of longer intervals. This variability is clearly not due to scale errors because it appears in the estimates of groups of different ages for whom the interval of 6 months would be located on a different part of the scale. For the same reason the logarithmic relationship for intervals up to about 6 months cannot be attributed to scale effects.

Table 2. *Mean estimates (and standard errors) of intervals (in inches)*

(Scale 10 in. = 'life'.)

Group of subjects	...	(i)	(ii)	(iii)	(iv)	(v)				
Mean age (years)	...	8·42	10·83	13·5	19 +	20–29	30–39	40–49	50–59	60–74
Number	...	25	23	22	30	19	31	33	40	7
Interval: Actual	Interval: As fraction of year									
1 day	0·003	—	—	0·9 (0·13)	—	0·46 (0·17)	0·35 (0·10)	0·46 (0·09)	0·56 (0·08)	0·73 (0·38)
1 wk.	0·019	2·0 (0·27)	1·9 (0·30)	1·8 (0·32)	0·7 (0·18)	—	—	—	—	—
3 wk.	0·058	1·9 (0·34)	—	—	—	1·1 (0·26)	1·0 (0·19)	0·71 (0·09)	1·3 (0·07)	1·2 (0·29)
8 wk.	0·154	—	3·1 (0·57)	3·1 (0·61)	—	—	—	—	—	—
10 wk.	0·19	3·6 (0·22)	—	—	1·3 (0·19)	—	—	—	—	—
4 mnth.	0·33	5·9 (0·50)	—	—	2·0 (0·19)	1·2 (0·23)	1·3 (0·24)	1·1 (0·13)	1·1 (0·17)	1·2 (0·24)
6 mnth.	0·50	—	4·1 (0·54)	—	—	2·4 (0·33)	2·1 (0·29)	2·2 (0·52)	1·8 (0·45)	2·0 (0·33)
7 mnth.	0·58	5·9 (0·52)	4·9 (0·53)	4·0 (0·50)	—	—	—	—	—	—
9 mnth.	0·75	—	—	—	—	2·1 (0·34)	2·0 (0·58)	2·4 (0·56)	2·2 (0·55)	3·7 (0·72)
12 mnth.	1·0	—	4·5 (0·38)	4·5 (0·52)	2·6 (0·22)	—	—	—	—	—
1 yr. 3 wk.	1·058	7·3 (0·41)	4·3 (0·34)	—	—	3·0 (0·32)	2·7 (0·76)	3·3 (0·57)	3·3 (0·36)	4·0 (0·57)
3 yr.	3·0	5·5 (0·40)	—	—	—	4·6 (0·49)	2·95 (0·50)	4·2 (0·65)	3·6 (0·27)	3·5 (0·46)
4 yr. 5 mnth.	4·45	—	7·5 (0·27)	5·9 (0·45)	—	—	—	—	—	—
7 yr. 6 mnth.	7·5	—	—	—	4·4 (0·44)	—	—	—	—	—
8 yr.	8·0	—	—	—	—	4·8 (0·43)	4·1 (0·44)	3·8 (0·60)	2·6 (0·44)	5·4 (0·63)
8 yr. 6 mnth.	8·5	—	—	8·5 (0·39)	—	—	—	—	—	—
13 yr. 6 mnth.	13·5	—	—	—	7·2 (0·26)	6·7 (0·39)	5·1 (0·32)	4·7 (0·50)	4·5 (0·62)	5·5 (0·69)
14 yr.	14	—	—	—	8·0 (0·23)	—	—	—	—	—
27 yr.	27	—	—	—	—	—	9·0 (0·18)	8·0 (0·17)	7·3 (0·27)	6·9 (0·44)
35 yr.	35	—	—	—	—	—	—	9·0 (0·02)	7·4 (0·18)	6·9 (0·50)

A further check on the possibility of the logarithmic effect being due to the scale was made by obtaining estimates (by the 13·5-year age-group) on a line representing 1 year from now. These estimates were compared with the corresponding estimates made on the scale representing life. They were found to yield a good fit to the corresponding part of the curve fitted to the estimates made on the life scale. The comparison is shown in Table 3.

Table 3. *Mean estimates of intervals* (*in inches*)

Interval		Group (iii) (Scale, 10 in. = 'life')	Group (vi) (Scale, 10 in. = 1 yr.)
Actual	As fraction of year		
1 day	0·0027	1·548*	1·24
3 days	0·008	—	2·46
1 wk.	0·019	3·09	2·89
4 wk.	0·076	—	5·31
5 wk.	0·096	—	4·65
8 wk.	0·154	5·33	—
3 mnth.	0·25	—	6·55
7 mnth.	0·58	6·88	6·66

* Figures in this column are multiplied by a constant (1·72).

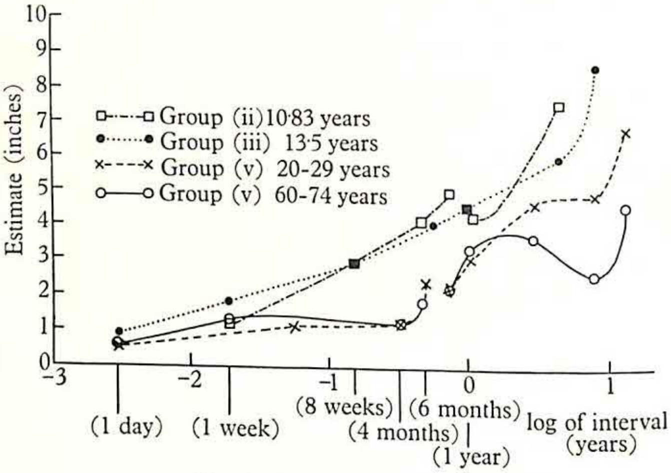


Fig. 1. Means of estimates.

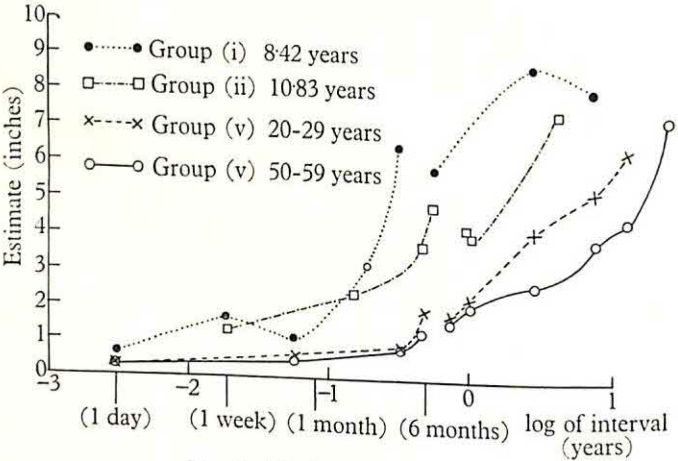


Fig. 2. Medians of estimates.

In Figs. 1 and 2 the curve representing the relationship between the intervals and the estimates shows a discontinuity at a constant point on the time scale (Y-axis), irrespective of where the point falls on the estimate scale (X-axis). This discontinuity thus appears to be a function of the time interval and not of the estimate scale. The discontinuity cannot be attributed to the actual delimiting events chosen because these varied from group to group. Furthermore, the experiments extended over a period of 3 months so that the same event delimited different intervals for different subjects and the same interval was delimited by different events.

The suggestion that two types of judgement respectively are made for intervals up to about 4–6 months and beyond 1 year is supported by (i) the bimodality of the distribution of estimates in the intervening zone, and (ii) the increased scatter in this zone. Coefficients of variation are plotted to illustrate the variation in scatter and they decrease fairly smoothly with more remote intervals, except for the intervening zone where there is a pronounced departure in the form of a hump. These coefficients are shown in Fig. 3.

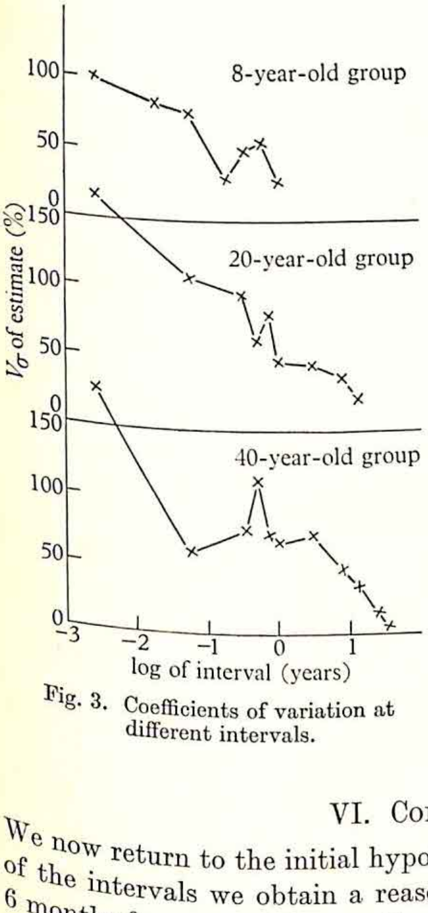


Fig. 3. Coefficients of variation at different intervals.

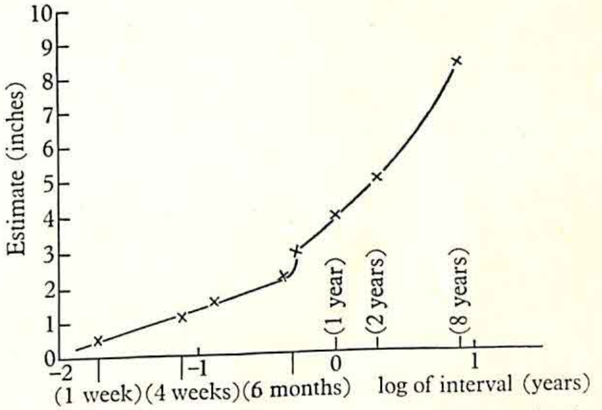


Fig. 4. Group (vii). Scale, 12 in. = 10 years. Mean estimates.

VI. CONSIDERATION OF HYPOTHESIS

We now return to the initial hypothesis. If we plot the estimates against the logarithm of the intervals we obtain a reasonable fit to a straight line for intervals up to about 6 months from now. We may therefore conclude that the expectation, based on Weber's law, that a logarithmic increase in chronological intervals should yield a linear series of estimates of these intervals, is supported by our data only for past intervals up to about 6 months.

It is of course possible that the above effects may be accounted for by other hypotheses with which we are not here primarily concerned. One such hypothesis might be that the seeming length of an interval may be a function of the retention of activity during the interval and, as retention diminishes with time, there might be a corresponding contraction of judgement of time intervals.

VII. SUMMARY

We have attempted to extend the quantitative study of psychological time to intervals of months and years. The method employed is based on the assumption that (*a*) one's total past life is a unity which can be used as a standard of comparison, and (*b*) that estimates of past intervals can be represented spatially.

Our results suggest that, at any given age, estimates of intervals since given past events may be interpreted in accordance with the Weber principle up to a period of about 6 months. Beyond this period our method appears to yield a linear relationship only. This may be because people are unable to make direct judgements of intervals beyond 6 months, such intervals being calculated by them on a chronological basis.

REFERENCES

COHEN, J., HANSEL, C. E. M. & SYLVESTER, J. D. (1953). A new phenomenon in time judgment. *Nature, Lond.*, CLXXII, 901.

HOAGLAND, H. (1935). *Pacemakers in Relation to Aspects of Behaviour*. New York: Macmillan. (And in many more recent papers.)

JAMES, W. (1901). *Principles of Psychology*, I, p. 625. London: Macmillan.

PIÉRON, H. (1952). *The Sensations*, pp. 289–305. Transl. by M. H. Pirenne and B. C. Abbott. London: Muller.

STURT, M. (1925). *The Psychology of Time*. London: Kegan Paul.

WOODROW, H. (1951). Time perception. In S. S. Stevens (ed.), *Handbook of Experimental Psychology*, pp. 1224–36. New York: Wiley.

(*Manuscript received* 1 *December* 1953)

SOME PROBLEMS OF ANOMALOUS COLOUR VISION

By R. W. PICKFORD

Psychology Department, Glasgow University

I. *Introduction* (p. 115). II. *The anomalous trichromat* (pp. 115–116). III. *The Rayleigh Equation* (pp. 116–117). IV. *Types of red and green anomaly* (pp. 117–118). V. *Terminological problems* (pp.118–120). VI. *Statistical problems* (pp. 120–121). VII. *Problems of heredity* (pp. 121–123). VIII. *Modes of colour vision variation* (p. 123). IX. *The trichromatic theory* (pp. 123–124). X. *Bipolar and zone theories* (pp. 124–126). XI. *Conclusions and Summary* (pp. 126–127). *References* (pp. 127–128).

I. INTRODUCTION

This paper is an attempt to explain anomalous colour vision in a simple way and to meet as far as possible various difficulties and problems concerning it. These problems are complex and by no means settled at present, but many people are interested in them, and both from practical and theoretical points of view they are important. They will be treated in a general way here, the principal aims being to set forth, to co-ordinate and as far as possible to harmonize various contributions to the whole subject, but in the space available no attempt can be made to give its complete history or to deal with all the problems in full detail.

II. THE ANOMALOUS TRICHROMAT

Owing to the work of Helmholtz and many others, especially of Wright (1946), it may be taken as fact that those with normal colour vision can match in hue all colours in the spectrum, and, of course, also those outside it, with appropriate mixtures of three suitably chosen homogeneous lights, namely red, green and blue. If we allow the hue being matched to be desaturated with one of the matching hues, then satisfactory matches for saturation as well as for hue can be made. This has led to much discussion, because it calls for the use of 'negative' quantities of the matching colours in certain combinations. However, these negatives are not large, and might be satisfactorily accounted for simply as measurements of the degrees of dilution or desaturation which are produced by the over-lapping of the responses or functions of three primary types of process or receptor, none of which can be excited without one or both of the others being also excited at the same time, and therefore the negatives do not necessarily contradict a strictly trichromatic theory. In principle only three variables are being manipulated. At the same time, of course, there might be more primaries, and the negatives might be due to still more complex overlapping and/or to the presence of additional primary receptors, functions or processes.

The trichromatic theory, however, assumes the smallest practicable number of kinds of receptors or processes, namely three, but the facts of trichromatic colour mixing could equally well be satisfied by a theory that there were more, say four, with Hering and Houstoun (1932), or seven, with Edridge-Green (1891), and Hartridge (1948). The trichromatic theory states that, however many kinds of receptor or function may be at work, their action must be such that it can be represented as if there were only three, though it must be added that some recent work by Hurvich & Jameson (1951), Pickford (1951), and Piéron (1952), tends to confirm many long-standing suspicions that there must be independent receptors contributing to brightness.

When adequate experiments have been done on sufficiently large numbers of people it is not impossible that a very small proportion will be found who require more than three primaries to match all colours. A small proportion of people appear to be dichromats, namely, less than 1·5% of men, of whom about 1% are protanopes and about 0·5% are deuteranopes, and these proportions are far smaller among women. It is not easy to be sure that a given defective is absolutely dichromatic, because of difficulties in testing, and because the result of a test depends partly on the area of the retina stimulated and the intensity of the stimulus, but to the best of our knowledge the small proportions mentioned are dichromatic. Of the remainder of the population of Britain, about 92·5% of men and 99·5% of women are normal trichromats, but there is a small amount of variation among them which is not detected at all by ordinary 'colour blindness' tests (Pickford, 1949*a*, *b*, 1950, 1951). The remaining 6%, approximately, of men, are anomalous trichromats, and they fail on the Ishihara Test, like dichromats.

III. THE RAYLEIGH EQUATION

Rayleigh (1881) described an experiment in which subjects were asked to match a variable mixture of red and green lights against a standard yellow. Lithium red, thallium green and sodium yellow have become the standard colours to be used for this purpose. He described three ways in which the experiment could be done, and he found that among twenty-three male observers five, three of whom were his brothers-in-law, required more green while two required more red than the majority. It was to these exceptional subjects, who were defective without being dichromats, that the term 'anomalous trichromat' was first applied. One apparatus which he used was subsequently employed by Schuster and by Edridge-Green (1920), both of whom tested larger numbers of subjects. Since this time many workers have used the anomaloscope, as the apparatus has come to be called, and in Houstoun's (1932) form of it the variable proportions of the red-green mixture were controlled by nicol prisms, while in Nagel's, which has become an established pattern (Lohmann, 1913; Willis & Farnsworth, 1952), the mixture is controlled by varying the widths of two slits in inverse proportions. Collins (1929) and subsequently Pickford (1951) showed that similar results could be obtained with rotating paper disks. Since essential characteristics of anomalous trichromats always concern the relative sensitivity to red and green in matching a yellow light, the Rayleigh Equation has become the most satisfactory test for their defects. Several types of anomaloscope have been devised (Willis & Farnsworth, 1952; Crawford, 1951*a*, *b*; Pickford, 1951).

Much research has been devoted to the colour vision of anomalous trichromats, but it appears that their characteristics are still to some extent in doubt, and their relative frequencies in various populations tested have not been sufficiently studied. When Holmgren's Wools test was the usual test for colour blindness, the total number of defectives among men appeared to be about 5%. With the use of Stilling's Tables and the Ishihara and other pseudo-isochromatic plates, the number rose to about 7·5% or more. These plates fail all the anomalous trichromats as well as the dichromats, but the true proportion of anomalous subjects must be settled by the use of the Rayleigh Equation with some form of anomaloscope, because pseudo-isochromatic plates do not sufficiently discriminate between the types (Chapanis, 1949*b*; Hardy, Rand & Rittler, 1945*a*–*c*, 1949; Pickford, 1949*b*, 1950, 1951). The so-called 'exaggerated simultaneous contrast' of the anomalous is probably an effect of colour naming, of perceiving rather than of sensation itself, or of judgement. It is not a reliable guide (Chapanis, 1949*a*; Pickford, 1951), and cannot be employed to diagnose the anomalous as suggested by Lohmann (1913) in his technique for using the Nagel card-test. The importance of adequate tests has already been emphasized (M.R.C. (1933); Riddell *et al.*, 1949; Pickford 1951; Wright, 1953) and will not be discussed further here.

IV. TYPES OF RED AND GREEN ANOMALY

Houstoun (1932) pointed out that there was a difficulty in the distribution of red and green deviations in the Rayleigh Equation. Some deviations fall within the normal curve and others fall outside it. He said that it would be best to reserve the term 'anomalous' for those subjects whose deviations fall right outside the normal curve. If this is done, the others may be called red and green deviants if they have deviations of about two to three times sigma (Pickford, 1949*a*, 1951). Houstoun also pointed out, on the basis of another set of experiments, that there is a proportion of subjects whose colour discrimination is good but different from that of the majority. He called these 'colour different', and possibly they are to be identified with the strictly anomalous subjects, whose matching ranges in the Rayleigh Equation are not at all or not much increased, but their deviations are of the order of six or seven times sigma, which is very far outside the normal curve.

It is increasingly clear that the concentration of attention upon deviation in the Rayleigh Equation, and the desire to give each subject a definite 'quotient' of red to green required to match the yellow standard, has to some extent obscured the facts. Walls & Mathews (1952) express clearly the views adopted by many workers, who are supporters of the trichromatic theory, namely, that all types of anomalous subjects must reject the normal Rayleigh Equation, and that with increasing defect, increasing deviation either to the red side for the protanomalous subjects, or to the green side for the deuteranomalous, leads through extreme anomaly, though not necessarily in a continuous way, to complete dichromasy. Nothing could be further from the truth.

An anomaloscope employing monochromatic colour filters, was used to measure systematically by an efficient psycho-physical technique the subject's average choice of red-green mixtures to match the yellow, and also the complete ranges of those matches, on a large number of subjects chosen at random and on many defectives, using several forms of test (Pickford, 1949*a*, 1951). It was apparent, as Edridge-Green (1920) and Franceschetti (1928) had observed before, that subjects differ not only in the deviations of their average red-green mixtures to one side or the other, but also in the magnitudes of their matching ranges. The relationship between deviation and matching range is complex and variable. Sloan (1950) and Crawford (1951*a*, *b*) have both described or mentioned the difference between the strictly anomalous who have normal or almost normal matching ranges but big deviations, and other incomplete trichromats who have greatly exaggerated red-green matching ranges either with or without deviations of six times sigma or more. The increases of matching range in question are very great, often as much as fifteen or twenty times the normal, and amount to 'colour blindness', though the subjects in question are not dichromats. Further work has been done along the same lines by Willis & Farnsworth (1952).

It is clear now that some major red-green defectives simply make a very abnormal Rayleigh Equation, and may be called truly anomalous or colour different. They fail unequivocally on the Ishihara Test. Others have greatly exaggerated red-green matching ranges when the yellow standard is kept at the same brightness for them as the matching mixture, but not as great as those of the true dichromats, who can match yellow with red or green or with any mixture of these colours, provided again that the mixture is of the same brightness for them as the yellow standard. The partial trichromats with exaggerated matching ranges may have strong deviations to the red or green side as well, but some do not. They also fail on the Ishihara Test.

The problem of the nature of the relation between matching range and deviation is complex. All those who are complete dichromats, of course, have no measurable deviation, but those who are almost dichromatic may have mid-points or deviations nearer the normal in the Rayleigh Equation than those who are anomalous trichromats and have small ranges and big deviations. This is probably difficult to understand on the trichromatic theory, but on the bipolar theory of Houstoun it is perfectly clear. The anomalous in the narrow sense do not form a series continuously graduated in frequency and deviation between trichromats and dichromats, nor are the deviations of those partial trichromats who have exaggerated matching ranges graduated evenly in this way either. Some subjects who have a lesser degree of trichromasy have relatively smaller deviations, and vice versa. This is also easier to understand on Houstoun's than on the trichromatic theory.

In addition there are two further subgroupings which are clear. First, among the 'red different' or protanomalous in the narrow sense, some have the red end of the spectrum, shortened or darkened to a statistically significant extent, while for others it is of normal length or brightness. Secondly, in the group of partial dichromats or moderately dichromatic subjects of the deuteranope class, who have exaggerated red-green matching ranges in the Rayleigh Equation, some have characteristically marked green deviations and reject the normal match, while others do not. It is clear, therefore, that deviation in the Rayleigh Equation is large when we are dealing with the strictly anomalous subjects, but becomes less and less as matching range increases, though these changes sometimes occur in a stepwise manner, until the condition of dichromasy is reached, and then deviation is zero and matching range infinite. This is very different from what has been assumed by some workers on the basis of the trichromatic theory (Walls & Mathews, 1952). Their assumption is that deviation must increase continuously until dichromasy is reached.

V. TERMINOLOGICAL PROBLEMS

The terminological difficulties are considerable. The best plan is to begin by confining the terms 'protanomalous' (PA), 'red anomalous' or 'red different', on the one hand, and 'deuteranomalous' (DA), 'green anomalous' or 'green different', on the other, to those partial trichromats who have good red-green discrimination, namely normal or almost normal matching ranges, but who have big deviations of the order of six or seven times sigma in the Rayleigh Equation. Then the terms 'extreme protanomalous' (EPA) or 'moderate protanope', and 'extreme deuteranomalous' (EDA) or 'moderate deuteranope', may be used for those partial trichromats who have exaggerated red-green matching ranges with or without large deviations. Those mentioned previously who are in this group but reject the normal equation may be called 'deviant deuteranopes', and are probably the same as the 'anomalous deuteranopes' of de Vries (1949). Similar subjects are occasionally met with among moderate protanopes, but are very infrequent and may be overlooked at present.

A considerable difficulty is created by the fact that the EDA group, as defined above, and to a lesser extent the EPA group, is rather too variable for all its members to be called 'extreme anomalous', after we exclude the deviant deuteranopes. It is true that some EDA and EPA subjects are almost dichromatic, and, of course, have very small mean deviations. They sometimes have mid-points on the 'wrong' side of the normal match. It is reasonable to call such subjects and the deviant deuteranopes 'extreme deuteranomalous' or 'deuteranomalous', but there is a proportion who have much smaller matching ranges, though still perhaps ten times as great as those of the average normal subject, and also rather small deviations or none at all. These subjects are clearly not to be classed as DA (or as PA if they are found in the protanope series), but their defects are much less than those of the obviously extreme anomalous. In spite of the troubles which might be caused by multiplication of terms, it seems that there would be an advantage in speaking of 'lesser' and 'greater' defectives in the ordinary EDA group at least, and possibly in the EPA group as well, though they are statistically continuous.

The terms 'protanope' (P), and 'deuteranope' (D), may be reserved for true dichromats, excluding 'tritanopic' dichromats, who are undoubtedly very rare (Wright, 1952). Nevertheless, the distinctions between the most extreme of the EPA and the P subjects on the one hand, and between the most extreme of the EDA and the D on the other, will not always be easy to make.

Table 1. *Major red-green defectives, totally sex-linked in heredity*

	Incomplete trichromats		Complete dichromats
Protans	1. Protanomalous, red anomalous or red different (PA): (*a*) with red of normal brightness; (*b*) with darkened red	2. Extreme protanomalous or moderate protanopes (EPA), not subdivided here	3. Complete protanopes, or red blind (P)
Deutans	4. Deuteranomalous, green anomalous or green different (DA)	5. Extreme deuteranomalous or moderate deuteranopes (EDA): (*a*) deviant; (*b*) ordinary (lesser and greater)	6. Complete deuteranopes, or green blind (D)

A criterion for the distinctions between PA and EPA, and between DA and EDA, which might be thought important, is that the PA and DA might confidently reject the normal match in the Rayleigh Equation, always provided that brightnesses are equated for the subject in question at that point. As already mentioned, however, all deviant deuteranopes and a few EPA subjects also reject it, and so this criterion cannot be used. The distinctions between PA and EPA and between DA and EDA therefore turn mainly on the size of the matching range, which is very small in PA and DA subjects but large or very large in EPA and EDA. The PA and DA subjects form homogeneous groups with large deviations and small matching ranges, whereas the EPA and EDA subjects always have large or very large matching ranges, and they are somewhat heterogeneous groups.

The terms 'protan', for all protanopic subjects, and 'deutan' for all deuteranopic subjects, introduced by Farnsworth, will be convenient, and no difficulties will be raised by employing them in these ways, as they have been used by Sloan (1950) and Chapanis (1951), and by Willis & Farnsworth (1952).

Table 1 shows the classification of major defectives discussed in this article, and may help to clarify the terminological difficulties. Minor red-green and yellow-blue defectives, who may be called deviant and colour weak trichromats, and who pass the Ishihara Test with a small number of errors or no errors at all, and major blue-yellow defectives, have been omitted.

VI. STATISTICAL PROBLEMS

The question arises in exactly what ways the so-called anomalous trichromats of various types are intermediate between normal subjects and extreme dichromats in the statistical sense of relative frequency. This problem can be decided only by painstaking study of large and random samples of various populations, and it is well-known to psychologists that the problems of obtaining representative random samples of populations are very considerable. Since it is already known that the frequencies of different kinds of colour vision defectives vary from one racial group and from one geographical group to another (Pickford, 1951), it is not satisfactory to assume until it has been proved, that in Sweden, Norway, Germany, France, Britain or the U.S.A., for example, or among different racial stocks in the same countries, the same types of defect will be found or that they will occur in the same proportions. Fortunately in colour vision we do not have to face the difficulty to be met in the study of the frequencies of left-handedness, for instance, which looks like a Mendelian recessive, that it is greatly modified in its manifestations by training and indirect environmental influences. We do, however, have to face what might be called the creation of types by the methods or weaknesses of tests and testing techniques. The lack of validity and reliability of the Ishihara and other tests, and the tendency of many workers to predict and even to assume types of defectives in accordance with one or other favourite theory, and to base their conclusions on too small and non-random samples of the populations tested, have caused many difficulties and confusions.

When the different types of so-called anomalous trichromats are studied in such a way as to group together those who are alike, it becomes clear that the colour different or truly anomalous form two statistically distinct groups (PA and DA), which would not become continuous, either with the normal population or with each other, however many subjects were tested (Pickford, 1951). Even if some intermediates between the normals, red anomalous and green anomalous, were found, there would never be enough of them to fill the gaps from the statistical point of view. The group of red anomalous or red different subjects (PA) is again subdivided in a statistically decisive way into those with and those without what is generally regarded as the special characteristic of protanopes, namely the darkening of the red end of the spectrum. There is, however, in respect of the distribution of greatly increased matching ranges, no very clear statistical distinction between the group of moderate protanopes or extreme protanomalous subjects (EPA) and the true protanopic dichromats, or between the group of moderate deuteranopes or extreme deuteranomalous subjects (EDA) and the true deuteranopic dichromats. The difference between deviant or anomalous deuteranopes and ordinary moderate deuteranopes looks as though it might be statistically significant if a large enough sample were tested. The distinction between all protans and all deutans, excluding the red anomalous without darkened red, is clear in terms of the shortening or darkening of the red end of the spectrum which occurs in the protanopic group, and there are no true intermediates of the kind created by the inefficiency of the Ishihara Test. The groups do not run together statistically.

For these reasons the distinctions between the colour different or truly anomalous and the whole groups of partial and complete dichromats of the types which have very large red-green matching ranges with or without big deviations, are clearer and more decisive than the distinctions between the moderate and complete protanopes or the moderate and complete deuteranopes, though the moderate deuteranopes must be divided into deviant and ordinary classes as explained above. It must be said that conclusions should be based on more extensive researches, using larger numbers of subjects chosen on a satisfactorily random and representative basis and also more tests all carried out on every individual. Willis & Farnsworth (1952) have made a valuable start in this direction, but they did not use a truly random sample. Unfortunately they have equated the mid-points and end-points, which are not physically or psycho-physically comparable, of six different types of anomaloscope test.

Table 2. *Frequencies of red-green major defectives in a group of* 142 *taken consecutively*

	PA		EPA (Not subdivided here)			P	Totals
	Red norm.	Red dark					
Protans	4	10	14			16	44
			EDA				
				Ordinary			
	DA		Deviant	Less	Great	D	
Deutans	52		17	14	10	5	98

Table 2 shows the proportions of different types of red-green major defectives in a group mainly tested in Glasgow all of whom failed the Ishihara Test outright and were also tested with a form of anomaloscope using monochromatic gelatine colour filters (Pickford, 1951, with more material). They were taken in the order in which they were found, and, although a small number were relatives, they represent an approximately random sample. Nineteen were women: 1 PA with normal and 3 with darkened red; 1 EPA; 12 DA; and 3 EDA of the deviant type. Probably not more than 10 women at the most would be expected by chance in a group of 142 major defectives. Ten EDA subjects had mean deviations somewhat to the red side of the normal mid-matching point, and two EPA slightly to the green side. It will be observed that among protans the number of individuals increases in each type as the defects increase, whereas among deutans the number decreases. Although deutans are more than twice as frequent as protans, there are three times as many protanopic as deuteranopic dichromats.

VII. PROBLEMS OF HEREDITY

Owing to the work of Waaler (1927) and others, it became clear that the defects of protanopes and deuteranopes were inherited by separate Mendelian recessive mechanisms, although both were totally sex-linked as realized before, and they are now regarded as dependent on genes having separate loci on the X-chromosome. Franceschetti (1928, 1949) added the so-called anomalous and extreme anomalous (PA and EPA, and DA and EDA), and pointed out that there must be two series of multiple allelomorphs, in each of which the gene for complete dichromasy was the most recessive, the normal gene the most dominant and the other genes were intermediate in order of the severity of the defects they produced. Many researches by other workers have supported this theory and it has been widely accepted. Walls & Mathews (1952) have suggested that true intermediates between the two series may occur, as for instance, when a man has protanope and deuteranope genes together, but of course in different loci of his X-chromosome.

There are, however, in all probability further complexities, and there are certain pedigrees suggesting strongly that 'deviant deuteranopes' or 'anomalous deuteranopes', and red anomalous or protanomalous subjects with red of normal brightness are due to two additional allelomorphic genes, giving one normal and four rather than three defective genes in each series (Pickford, 1951). The fact that extreme protanomalous and extreme deuteranomalous subjects are very diverse and statistically indistinct from protanopic and deuteranopic dichromats respectively, makes it seem that there could be even more genes in action or that varying 'penetrance' and 'expressivity' could make marked differences to their effects. A further problem arises because there are certain defectives who are not clearly distinct from the extremes of the curves showing the distributions of normal deviations and matching ranges in the Rayleigh Equation, but whose defects are inherited true to type and, at least sometimes, definitely on the totally sex-linked pattern, excluding the normal heterozygotes for the familiar major defects. These are 'extreme deviants', and those who are exceptionally 'colour weak' in red and green, having fairly large matching ranges. Their defects are probably due to lesser defective genes, dominant to the genes for PA and DA but still recessive to the genes for normal colour vision in each series (Pickford, 1953*a*). It is likely, therefore, that in so far as the principal forms of red-green defect and their most marked minor forms are concerned, we must think of not less than five defective allelomorphs in each series, all, of course, totally sex-linked and some rather variable.

It is also very likely that, when women heterozygotes showing slight effects due to the familiar defective genes are excluded, there will be a number of variations of normal colour vision, inherited either by additive genes, or by non-additive genes partially sex-linked or not sex-linked at all, giving on the average equal proportions of defectives in the two sexes (Pickford, 1951). These observations apply also to yellow-blue defects, (Pickford 1953*a*), though there may be more extreme forms of tritanopia inherited in other ways.

Various researches have suggested that major defects are incompletely recessive. Careful tests of normal heterozygotes show that there is a strong tendency for diminished forms of the types of major defect found among their male and female relatives to be present in them (Pickford, 1951). Normal heterozygotes of the red anomalous or protanomalous class, and of the green anomalous or deuteranomalous class, tend to show small red or green deviations accordingly. Protanopes, moderate protanopes and extreme protanomalous subjects, and deuteranopes, moderate deuteranopes and extreme deuteranomalous subjects, all tend to have normal heterozygote relatives who show large red-green thresholds or matching ranges. Deviant or anomalous deuteranopes tend to have normal heterozygote relatives who have marked green deviations in addition. The problem of the colour vision types of the abnormal heterozygotes, namely those who inherit two or more defective genes, is too complex to be discussed here, but PA may sometimes be only partially dominant to EPA (Pickford, 1953*b*).

The generalization may be made, for both major and minor forms of defect, and whether they are totally sex-linked or not, that the more that matching range is the main feature, as with extreme anomalies and dichromasy, the more recessive the defect in the Mendelian sense, and the more that deviation is the main feature, the more dominant the defect.

VIII. MODES OF COLOUR VISION VARIATION

There are three principal ways in which colour vision differences may be viewed from the sensory aspect, in comparison with normal colour vision. These modes of variation include small differences, classed as minor defects, as well as large differences or major defects, such as anomalous colour vision and colour blindness. The three modes of variation, which sometimes occur together, are as follows.

(1) Increases in the range of matching for the discrimination of red and green mixture with a yellow standard, and of a yellow and blue mixture with a colourless grey. Small increases lead to colour weakness; large increases to partial and eventually to complete colour blindness for the two colours in question. Yellow and blue defects will be omitted from any further discussion.

(2) Deviations from the normal mid-matching point in the Rayleigh Equation. Small deviations up to about three times sigma cause deviant colour vision. Big deviations, of about six or seven times sigma give colour differences or truly anomalous colour vision, and deviations to the limits of the equation on either side would cause dichromasy because then the standard yellow could be matched with both variables in all mixtures throughout the scale.

(3) Darkening of one of the hues called primaries, or shortening of the red or the violet end of the spectrum. In red-green blindness this shortening or darkening affects the red end of the spectrum and is found exclusively in protanopes and in the majority of protanomalous subjects. Shortening or darkening of the violet end of the spectrum is found in a proportion of yellow-blue defectives.

These three modes of variation at the sensory level are not necessarily equivalents of three different functions at the physiological levels (assuming that the sensory level is the level of conscious experience), but they are the end products of complex physiological processes which must be unravelled for the purposes of colour vision theories.

IX. THE TRICHROMATIC THEORY

On the trichromatic theory we should expect at least two kinds of colour vision defects: (1) those due to losses of receptors or perhaps of pathways ('loss' defects); (2) those due to lack or breakdown of the difference between two kinds of receptors or pathways without actual loss ('collapse' defects). Loss defects would lead towards deviations, colour difference and truly anomalous colour vision (PA and DA), and could, if extreme, produce dichromasy. It is not clear that they would produce increased matching ranges until the dichromatic point was reached. There would be a high correlation between degree of deviation and darkening of the part of the spectrum corresponding to the loss, and also with magnitude of defect apparent in daily life. However, the red end of the spectrum is not darkened in about half of the protanomalous subjects (PA), and it is apparently as much darkened in some of these and in extreme protanomalous subjects (EPA) as in many dichromatic protanopes. In other words, there is no correlation between magnitude of loss and darkening of red. This lends support to other evidence that brightness must be mediated at least in part by independent receptors, as suggested by Piéron (1952).

Since the brightness of red may vary independently of sensitivity to red, it is difficult to understand how brightness could be mediated mainly by a central relay system which was dependent on the integrity of the red receptors.

These facts also make us doubtful whether red difference or protanomaly can always be a loss defect. Extreme protanomaly or moderate protanopia would not be likely to be a pure loss defect either, because in it there is a great increase of the red-green matching range as well as a very variable degree of red deviation and loss of brightness of red. Complete protanopia, amounting to dichromasy, however, might be a pure loss defect on this theory, but it would be a useful experiment to find out whether the loss of brightness of the red end of the spectrum was the same in all dichromatic protanopes. On the other hand, the green different or strictly deuteranomalous (DA), who would be expected to be pure loss defectives because of their great deviations to the green side in the Rayleigh Equation, never show the loss of brightness of green which would be expected on the trichromatic theory. Since brightness of green is not reduced in extreme deuteranomalous subjects and dichromatic deuteranopes, it seems clear that they cannot be loss defectives, and this has been widely realized.

Collapse defects would lead towards increases of the red-green matching range, without darkening of either hue, and without changes or deviations of mid-points in the Rayleigh Equation, unless loss defects were combined with them, when the deviation would be towards the side of the loss in question. Then there would be a high correlation between magnitude of deviation and darkening of the appropriate colour. Hence it seems clear that the protanomalous with red of normal brightness can neither be pure collapse defective nor combine collapse and loss. Deuteranomalous subjects (DA) cannot be pure loss or collapse defectives, and, since moderate and deviant deuteranopes (EDA) all show variable degrees of green deviation uncorrelated with their increases of red-green matching range, it would appear that they cannot be pure collapse defectives either, but many of them, it seems, particularly the deviant or anomalous deuteranopes, might combine loss and collapse defects. This, however, is not possible, because none of them show darkening of green however great their green deviations, and a few, in the writer's experience, even have red deviations, just as a few protanomalous subjects have deviations of mid-points to the green side, combined with darkening of the red.

All these difficulties are felt in similar ways on a smaller scale in dealing with minor defectives, that is to say, those who have small increases of matching range and small deviations. It is unfortunate that such considerations shake our faith in the trichromatic theory very severely, and they led the writer to relinquish it altogether in the past, but it would be unscientific to be in any way dogmatic or to regard the matter as settled, though it does not help the trichromatic theory to say that the physiology of colour vision is so complex that anything might happen.

X. BIPOLAR AND ZONE THEORIES

On a modified form of the ingenious hypothesis due to Houstoun (1932), following Hering's theory, it would appear that the physiological processes underlying the sensations of red and green might be controlled by a bipolar system. There is much evidence pointing in this direction from other sources, and many parallel considerations apply to yellow and blue which might also be controlled in a bipolar manner. On such a system the difference between deviation and matching range defects could be explained much more efficiently and economically than on the trichromatic theory as usually understood. If yellow sensitivity did not vary independently of red and green it might be possible to think of a trichromatic theory in which red and green formed one bipolarity, and another bipolarity was formed between red and green on the one hand and blue on the other, after the fashion of Ladd-Franklin's scheme, but unfortunately this is not possible because yellow does vary independently.

On Houstoun's theory we may suppose that the red-green pair of receptors or of processes or functions have a point of change-over from the 'red' signal or response to the 'green' signal, whatever their physiological nature may be, and that the same may be true for yellow and blue, although of course, yellow is contributed partly by an additive process of red and green when we are using stimuli which do not correspond to complementary sensations. If all the red-green receptors or pairs of receptors or functions change over together and at the usual point, we have normal red-green discrimination and no deviation. If a number tend to change over at various points or irregularly, we tend to have increased red-green matching ranges, most extreme when all responses are in one direction alone, or half in each direction at the same time. When most of the responses tend to change over at a particular though exceptional point, we have deviation, colour difference or true anomaly. Loss defects would probably also have to be taken into account.

On this theory there would be an independent brightness receptor, though brightness is, as Hering supposed, contributed in part by hue responses as well. Major defectives excluded, it seems that losses of sensitivity to hue are correlated with decreases of brightness to a lesser extent for red and yellow than for blue and green (Pickford, 1951). Among protans (PA, EPA and P), however, loss of brightness of red is irregularly connected with losses of sensitivity to red in the ways described above.

Nobody will wish to deny that pigmentation of the transparent parts of the optical system may affect colour vision and brightness sensitivity and the relationship between them in certain ways, but it would be difficult to see how it could account for more than a small proportion of the variations reported. It is not impossible that a system such as that of Houstoun might be combined with a trichromatic system in some complex way such as that described by Talbot (1951), after Müller and Adams, or by Thomson (1952), after Müller, and brightness or whiteness might be controlled partly by a central relay system dependent on the hue receptors, but it is very unlikely to be completely dependent in this way. Some tests and experiments seem to stress the trichromatic nature of colour vision, while others emphasize its bipolar and tetrachromatic characteristics, but it is probably important for us to think that these need not be incompatible. It is probable, however, that the tetrachromatic theory of Houstoun would save a whole zone, on the zone theories of Müller, Adams, Talbot and Thomson, and this would be a considerable economy of hypothesis, though Nature does not necessarily work in the most economical way possible.

The form of Houstoun's theory suggested here would satisfy the requirements of the trichromatic aspects of colour vision, as Houstoun himself intended. It would give an adequate and easily intelligible account of all forms of red-green and yellow-blue defects and variations, and would make it clear, for example, how defects in yellow sensitivity are often connected with brightening of yellow, while blue defects are not necessarily accompanied by darkening of blue or violet. The theory would also explain the fact that many of the subjects classed as extreme protanomalous or deuteranomalous defectives show a chracteristic fluctuating kind of sensitivity in red-green discrimination. They call the red in the Rayleigh Equation sometimes 'green' when it has enough red in it for them to distinguish it from the yellow, and sometimes 'red'. While they call the green sometimes 'red' and sometimes 'green'. They may feel confused themselves, and these responses might be due to a fluctuating variability of the balanced response mechanism. Sometimes most of the receptors may give either a 'red' or a 'green' signal to the red or to the green stimulus, but irregularly. Whereas in other defectives who have the same degree of defect the responses may be consistent, always the 'red' signal for the red stimulus, and always the 'green' signal for the green stimulus, when it has enough of the appropriate light in it to be discriminated from the yellow.

It seems that a bipolar theory, perhaps along the lines of the Hering-Houstoun theory, or of another kind, or possibly a zone theory in which one zone is concerned with the bipolar aspects of colour vision, is necessary in order to understand the facts of anomalous colour vision adequately.

XI. CONCLUSIONS AND SUMMARY

1. The concept of anomalous trichromasy should be reserved for several forms of major red-green colour vision defect, and other terms, namely 'deviant' and 'colour-weak' should be used of those who have minor defects in red and green (and also in yellow and blue).

2. The Rayleigh Equation is the most satisfactory simple test for anomalous colour vision, if used in such a way as to measure the subject's entire matching range, from which his mean deviation can be obtained. Brightness equations must always be obtained before colour equations are accepted.

3. The main types of anomaly are red anomaly or protanomaly, PA, with and without darkened red; green anomaly or deuteranomaly, DA; extreme protanomaly and deuteranomaly, EPA and EDA. The latter are divided into ordinary and deviant groups, and both are distinguished from PA and DA respectively by the magnitude of the matching range rather than the deviation, though this is usually smaller. There are also probably lesser forms of protanomaly and deuteranomaly which are difficult to distinguish from the extreme forms of normal red and green deviation, but belong to the sex-linked series.

4. The frequencies of these types are such that in a random sample of 142 major defectives, although protans were about half as frequent as deutans, protanopic dichromats, P, were three times as frequent as deuteranopic dichromats, D, and the DA group were more than three times as frequent as the PA group.

5. In heredity there are two series of about five multiple allelomorphs, all totally sex-linked, in which deviation defects tend to be dominant and matching range defects recessive.

6. Anomalous colour vision is dangerous in all situations in which life and safety depend on red, yellow or amber and green signals, and adequate tests must be used for measuring it.

7. Thorough empirical studies of all colour vision tests should also be undertaken, using adequate statistical techniques and properly constituted random samples of populations.

8. A bipolar theory of colour vision, in one or another form, is required to account for the facts of anomalous colour vision, and also for minor variations of deviation and matching range.

References

Chapanis, A. (1949*a*). Simultaneous chromatic contrast in normal and abnormal color vision. *Amer. J. Psychol.* LXII, 526–39.

Chapanis, A. (1949*b*). Diagnosing types of color deficiency by means of pseudo-isochromatic tests. *J. Opt. Soc. Amer.* XXXIX, 242–9.

Chapanis, A. (1951). Vision. *Annu. Rev. Psychol.* II, 54–64.

Collins, M. (1929). The Rayleigh Equation with rotating disks. *Brit. J. Psychol.* XIX, 387–93.

Colour vision requirements in the Royal Navy (1933). *Spec. Rep. Ser. Med. Res. Coun., Lond.*, no. 185. London: H.M. Stationery Office.

Crawford, A. (1951*a*). Description of an anomaloscope. *Brit. J. Physiol. Optics*, VIII, no. 3, 173–5.

Crawford, A. (1951*b*). Rayleigh Equations on an anomaloscope. *Brit. J. Physiol. Optics*, VIII, no. 4, 225–34.

Edridge-Green, F. W. (1891). *Colour Blindness and Colour Perception.* London: Kegan Paul.

Edridge-Green, F. W. (1920). *The Physiology of Vision.* London: Hall.

Franceschetti, A. (1928). Die Bedeutung der Einstellungsbreite am Anomaloskop für die Diagnose der einzelnen Typen der Farbensinnstörungen, nebst Bemerkungen über ihren Vererbungsmodus. *Schweiz. med. Wschr.*, LVIII, 52, 1273.

Franceschetti, A. (1949). Première observation d'une fratrie issue de deux daltoniens de type different. *Bull. Acad. suisse Sci. méd.* V, no. 4, 227–32.

Hardy, L. H., Rand, G. & Rittler, M. C. (1945*a*) Tests for the detection and analysis of color-blindness. I. The Ishihara Test: an evaluation. *J. Opt. Soc. Amer.* XXXV, no. 4, 268–75.

Hardy, L. H., Rand, G. & Rittler, M. C. (1945*b*). Tests for the detection and analysis of color-blindness. II. The Ishihara Test: comparison of editions. *J. Opt. Soc. Amer.*, XXXV, no. 5, 350–6.

Hardy, L. H., Rand, G. & Rittler, M. C. (1945*c*). Tests for the detection and analysis of color-blindness. III. The Rabkin Test. *J. Opt. Soc. Amer.*, XXXV, no. 7, 481–91.

Hardy, L. H., Rand, G. & Rittler, M. C. (1949). The American Optical Company's Compilation of pseudo-isochromatic plates as a means of detecting and analyzing defective color vision. *J. Gen. Psychol.* XL, 103–17.

Hartridge, H. (1948). Recent advances in colour vision. *Advanc. Sci.*, V, 19, 243–53.

Houstoun, R. A. (1932). *Vision and Colour Vision.* London: Longmans, Green and Co.

Hurvich, L. M. & Jameson, D. (1951). A psychophysical study of white. I. Neutral adaptation. II. Neutral adaptation. Area and duration as variants. III. Adaptation as variant. *J. Opt. Soc. Amer.* XLI, 521–7, 528–36 and 787–801.

Lohmann, W. (1913). *Disturbances of the Visual Function.* London: Bale.

Pickford, R. W. (1949*a*). Individual differences in colour vision and their measurement. *J. Psychol.* XLVII, 153–202.

Pickford, R. W. (1949*b*). A study of the Ishihara Test for colour blindness. *Brit. J. Psychol.* XL, 71–80.

Pickford, R. W. (1950). An item analysis of the Ishihara Test. *Brit. J. Psychol.* XLI, 52–62.

Pickford, R. W. (1951). *Individual Differences in Colour Vision.* London: Routledge and Kegan Paul.

Pickford, R. W. (1953*a*). Inheritance of minor colour vision variations. *Nature, Lond.*, CLXXI, 1167.

Pickford, R. W. (1953*b*). A pedigree for three types of colour vision defect. *Nature, Lond.*, CLXXII, 1010.

Piéron, H. (1952). *The Sensations.* London: Müller.

Rayleigh, Lord (1881). Experiments on colour. *Nature, Lond.*, XXV, 64–6.

Riddell, W. J. B. *et al.* (1949). Discussion on colour vision in industry. *Proc. R. Soc. Med.* XLII, 145–50; *Sect. Ophth.* 1–6.

Sloan, L. L. (1950). Comparison of the Nagel anomaloscope and a dichroic filter anomaloscope. *J. Opt. Soc. Amer.* XL, 41–7.

Talbot, S. A. (1951). Recent concepts of retinal color mechanism. I. Contributions from psychophysics. II. Contributions from anatomy and physiology. *J. Opt. Soc. Amer.* XLI, 895–918 and 918–41.

Thomson, L. C. (1952). A review of ideas on colour perception. *Ophthal. Lit.* VI, no. 1, 3–40.

Vries, Hl. de (1949). The fundamental response curves of normal and abnormal dichromatic and trichromatic eyes. *Physica*, xvi, 367–80.

Waaler, G. H. M. (1927). Über die Erblichkeitsverhältnisse der verschiedenen Arten von angeborener Rotgrünblindheit. *Z. indukt. Abstamm.- u. Vererbungslehre*, xlv, no. 4, 297–333.

Walls, G. L. & Mathews, R. W. (1952). New means of studying color blindness and normal foveal color vision. *Univ. Calif. Publ. Psychol.* vii, no. 1, 1–172.

Willis, Marion P. & Farnsworth, D. (1952). Comparative evaluation of anomaloscopes. *Med. Res. Lab. Rep. No.* 190. Bureau of Medicine and Surgery, U.S. Navy Dept.

Wright, W. D. (1946). *Researches on Normal and Defective Colour Vision.* London: Kimpton.

Wright, W. D. (1952). The characteristics of tritanopia. *J. Opt. Soc. Amer.* xvii, 509–21.

Wright, W. D. (1953). Defective colour vision. *Brit. Med. Bull.* ix, no. 1, 36–40.

(*Manuscript received* 4 *September* 1953)

A NOTE ON EYSENCK'S ARTICLE 'THE EFFECTS OF PSYCHOTHERAPY: AN EVALUATION'

By LESTER LUBORSKY
The Menninger Foundation

In *The Journal of Consulting Psychology*, Eysenck (1952) surveyed reports on improvement of neurotic patients after psychotherapy and found that '...the results compared [well] with the best available estimates of recovery without benefit of such therapy. The figures fail to support the hypothesis that psychotherapy facilitates recovery from neurotic disorder. In view of the many difficulties attending such actuarial comparisons, no further conclusions could be derived from the data whose shortcomings highlight the necessity of properly planned and executed experimental studies into this important field.'

Dr Eysenck's 'experimental group' draws on the cases in twenty-four articles on the result of treatment by psychoanalytic and eclectic psychotherapeutic methods. The two sources of his 'control groups' are Landis's (1938) statistics on the consolidated amelioration rate of the state hospitals and Denker's (1946) survey of 500 consecutive disability insurance claims due to psychoneurosis treated by general practitioners. The main implicit argument runs that it is a matter of indifference whether psychotherapy is given or not since 'the figures' are indeterminate and so, it is explicitly stated, we should not insist on training clinical psychologists as psychotherapists! The latter point is a real *non sequitur*. If medical doctors had followed such advice their entire science would not have developed.

I am very much impressed with the negative results cited, based on the heavy array of statistics, and with the necessity for doing properly planned and executed research. Unfortunately, the defects of these studies are not presented fully or taken seriously enough. Some readers might jump to hasty conclusions that go beyond the principal fact: we do not have any studies which validly compare psychotherapeutic treatment results with the absence of such treatment.[1]

Specifically, I have two main criticisms and several minor ones. I do not believe Eysenck has an adequate control group nor that comparisons of groups can be made within the experimental group. Secondly, judgements of 'improvement' after treatment in these studies are very ambiguous.

To conclude as he does, Eysenck must assume patients do something they do not do: randomly self-select themselves to psychiatrists, general practitioners, and state hospitals. His 'treatment' versus 'non-treatment' groups are composed along very selective lines. Many of the 'insurance' group would probably never have visited the doctor if it were not required. As a whole the group is probably of higher social and economic level than other groups. Very likely the choice of a general practitioner rather than a psychiatrist to treat their psychoneurosis reflects a not-to-be-ignored difference in attitude to their illness. And, for all we know, different brands of psychiatrists may get self-selectively different samples of patients. Eysenck says we need better control groups, but fails to take his own statement seriously. Perhaps he is impressed by the size of his samples and forgets the lesson of the famous Literary Digest poll.

Pursuing his thesis, he also purports to show that different forms of treatment get different results. Some evidence is available on the obvious differences in the initial state of the patients who are given different forms of psychotherapy. At The Menninger Foundation the Psychotherapy Research Project is developing a scale for rating mental health.[1] Health ratings on this scale by psychotherapists and supervisors of psychotherapy confirmed our impression that the patients who start psychoanalysis begin on the average at a higher level than those who are given other forms of psychotherapy. Those who are hospitalized begin much lower than those who are given out-patient treatment. It is very likely, therefore, that when Eysenck reports that 44% improve with psychoanalysis, 66% with eclectic treatment, and 72% with custodial treatment he is talking about groups of patients who are grossly different.

---

[1] Eysenck neglects the *one* outstanding exception, Rogers & associates (1949, 1950, 1951).

At the very least what is needed for an adequate control group is elementary matching of groups of patients on level of mental health, symptoms, and duration of illness. Both control and experimental groups also should need treatment equally, but the former should need it and not get it for a definite period. Even this additional statement, 'and not get it', represents the beginning of what might be serious incomparability. The fact that a person tried to get treatment and did not get it can have varying effects: for some people it can pull them together, for others it can worsen their state. The only study I know (Rogers & associates 1949, 1950, 1951) that used an 'own-control' method in which the patient is retested after a period of waiting, confirms the fact that such waiting makes a difference. Further requirements for an adequate control group are described in a recent article by Hunt (1952).

About the second main point: the terms 'recovered, improved, slightly improved, unimproved' are useless for Eysenck's purposes. The terms say nothing about what the patient was like at the beginning and end of treatment; they can be and are applied to patients at the entire range of mental health. A schizophrenic patient can be called 'recovered'; so can a patient with a slight personality problem. Obviously the word 'recovered' is used differently in each case. An absolute scale avoids this pitfall of conventional improvement scales. For instance, as compared with patients in expressive psychotherapy, psychoanalytic patients began at a higher level of mental health. One would expect therefore they would have less room for improvement. Health ratings by psychotherapists and their supervisors of the Menninger Clinic patients' initial and later state showed this to be true.

Another reason for the discrepancies in 'improvement' judgements is the varying point of view of the judges as to what constitutes improvement or recovery. For example, analysts may be more strict about the kind of personality change they consider recovery —or their patients are more strict about the standards they wish to apply to themselves. It is likely that general practitioners would consider it 'improvement' if the patient simply went back to work, while for the psychoanalytic and eclectic therapists the judgement of 'improvement' would probably be based on personality change. Probably most patients in psychoanalysis were working or were able to work up to treatment, and continued through treatment and thereafter.

A few other minor points:

(*a*) Some of the elements of psychotherapy (e.g. support and suggestion) are present in a large number of the general medical practitioners' treatments. Also as Landis (1938) points out (in objecting himself to the use of the consolidated amelioration rate as a baseline for 'spontaneous' recovery) neurotics in state hospitals are given a variety of treatments, including some psychotherapy. And, as they are relatively unusual occupants of state hospitals, they probably get unusual treatment. At least, it is not entirely true that Eysenck was comparing a no-treatment with a treatment group.

---

[1] The scale consists of 100 points, where 100 represents ideal health and 0 total disintegration. Many intermediate points are defined and exemplified by thirty-two graded case descriptions. In making a rating the judge must balance off our seven criteria of health, consider the approximate position of his case within the sample case series, and then arrive at a single global rating. The scale is unpublished, but available to research personnel in mimeographed form.

(*b*) Eysenck includes under the category 'not improved' also 'died'. These should be taken out of his samples. Death, obviously, is not on any improvement-unimprovement continuum. Furthermore, the deaths in the samples are inconsistently included.

(*c*) Under the psychoanalytically treated group, Eysenck includes reports of results of treatment done from 1920 to 1930 in Fenichel's study (1930).[1] These data are older than others reported, the large number of cases over-weights the others in the subgroup, and a brief can be made that theory and technique have improved since those days. (But try and prove that one!)

In general, Eysenck is much impressed by the ease with which one's subjective feelings can be wrong. He quotes with approval the findings of the Michigan Project on the Selection of Clinical Psychologists (Kelly & Fiske, 1950) to the effect that the greater the *certainty* the poorer the *accuracy* of clinical judgement.

This was often but *not* consistently true—clinicians were sometimes both certain and accurate. (See later report, Kelly & Fiske, 1951.)

There are many difficulties in psychotherapy research. Clouding the facts and coming to wrong conclusions increases these difficulties. Clinical psychologists who have training and understanding in the psychotherapeutic process are in a good position to attack many tough problems involved in psychotherapy research; those who have withdrawn cannot help us to understand.

REFERENCES

DENKER, R. (1946). Results of treatment of psychoneuroses by the general practitioner. A follow-up study of 500 cases. *N.Y. St. J. Med.* XLVI, 2164–6.

EYSENCK, H. J. (1952). The effects of psychotherapy: An evaluation. *J. Cons. Psychol.* XVI, 319–24.

FENICHEL, O. (1930). Statistical report on therapeutic activities. In *Ten Years of the Berlin Psychoanalytic Institute*, 1920–1930 (ed. by the German Psychoanalytic Society). Vienna: International Psychoanalytic Association of Vienna.

HUNT, J. McV. (1952). Toward an integrated program of research on psychotherapy. *J. Cons. Psychol.* XVI, 237–46.

KELLY, E. L. & FISKE, D. W. (1950). The prediction of success in the VA training program in clinical psychology. *Amer. Psychol.* V, 395–406.

KELLY, E. L. & FISKE, D. W. (1951). *The Prediction of Performance in Clinical Psychology*. Ann Arbor, Mich.: University of Michigan Press.

LANDIS, C. (1938). Statistical evaluation of psychotherapeutic methods. In S. E. Hinsie (Ed.), pp. 155–65, *Concepts and Problems of Psychotherapy*. London: Heineman.

ROGERS, C. R. & ASSOCIATES (1949, 1950, 1951). *A Study of the Process and Outcomes of Client-Centered Therapy*. First, second, and third interim reports. Mimeographed, Counselling Center, University of Chicago.

---

[1] One member of the Psychotherapy Research Group, Dr Hellmuth Kaiser, formerly worked on the staff of the Berlin Psychoanalytic Institute from which Fenichel's report came. He recalls that many of the patients were treated in the Polyclinic by analysts who were treating their first cases. The judgements of improvement were made without any attempt at follow-up. Furthermore, the cases were screened for treatment very casually; many were completely unsuitable for psychoanalysis. The therapists' ratings may be too favourable or too unfavourable to psychoanalysis—their ratings help little in knowing the true state of affairs.

(*Manuscript received* 22 *October* 1953)

A REPLY TO LUBORSKY'S NOTE

By H. J. EYSENCK

Institute of Psychiatry, University of London

Luborsky appears to have two objections to my paper on 'The effects of psychotherapy'. The first one is that the conclusions do not follow from the data; the second one is that 'the main implicit argument', based on these conclusions, is a *non sequitur*. He seems to feel strongly on these points as he appears to accuse me of 'clouding the facts and coming to wrong conclusions'.

I find it difficult to answer his paper as I agree with most of what he says throughout. He is concerned with pointing out the shortcomings of the researches I have summarized, both from the point of view of selection of control and experimental groups, and from the point of view of the clinical judgements made of the improvement or otherwise of the patients. With these criticisms I agree; I have myself commented, in the article criticized, on the poor quality of the data 'whose shortcomings highlight the necessity of properly planned and executed experimental studies into this important field'. I am grateful to Luborsky for dotting the i's and crossing the t's, but I fail to see how this amounts in any sense to a criticism since it supports what I myself have written.

Basically, I think, Luborsky has misunderstood the logic of my approach. He seems to imagine that I tried to prove from the existing data that psychotherapy has no effect; this would be foolish (because you cannot prove a negative) as well as inadmissible (because of the poor quality of the evidence). What I tried to do, however, was something quite different: I tried to show that the available evidence, in spite of, or because of, its poor quality, fails to support the hypothesis that psychotherapy has any beneficial effects. I did not try to prove the null hypothesis, but merely examined the efforts of other investigators to disprove it. My conclusion was, and Luborsky does not attempt to deny its correctness, that these efforts ended in failure. As I point out in my conclusion: 'The figures fail to support the hypothesis that psychotherapy facilitates recovery from neurotic disorder. In view of the many difficulties attending such actuarial comparisons, no further conclusions could be derived.'

I do not see that Luborsky in any way disagrees with this statement; if he has found data to support the hypothesis of psychotherapeutic usefulness he certainly fails to mention them. He does say explicitly 'we do not have any studies which validly compare psychotherapeutic treatment results with the absence of such treatment'. I take this to be in complete agreement with my own conclusion that there is no evidence to show that psychotherapeutic treatment is successful. (Luborsky says in a footnote to the above sentence, 'Eysenck neglects the *one* outstanding exception'. This exception turns out to be a set of mimeographed interim reports of an experiment which, when it is completed and published, will undoubtedly be relevant to the problem under consideration.)

With regard to Luborsky's first criticism, then, we must conclude that it rests on a misunderstanding. Whether we are willing to accept the researches reviewed by me, bad as they are, as being of interest, or whether we reject them all, as Luborsky would seem to wish us do, we come to the same conclusion, namely, that *there is no scientific evidence* *regarding the beneficial effects of psychotherapy.* This is the point I wished to make and this is the point which Luborsky apparently agrees with.

What consequences follow from this demonstration? The views expressed in my paper were these: 'Until such facts as may be discovered in a process of rigorous analysis support the prevalent belief in the therapeutic effectiveness of psychological treatment, it seems premature to insist on the inclusion of training in such treatment in the curriculum of the clinical psychologist.'

Luborsky regards this as a *non sequitur*, i.e. he does not feel that this conclusion follows logically from the premises. I may have been at fault in not stating my major premise and not putting down in detail the intermediate steps of the argument. The main reason for this neglect was that they seemed to me self-evident; apparently I was wrong in this assumption. My main premise, then, is that psychology is a science; in other words, that it deals with observable facts and verifiable theories and that that which is not observed or verified does not form part of psychology. A clinical psychologist on this view would be a psychologist who applies approved scientific methods in the abnormal field and uses his scientific knowledge in dealing with the problems that arise. To apply methods having no proper rationale and to use treatments, the hypothesis of whose effectiveness is not supported by any empirical studies, does not appear to be the proper exercise of his scientific training, and it would seem to follow that such training should not include therapeutic procedures until their effectiveness had been proved beyond reasonable doubt. Lehner's comment seems apposite here, referring to 'therapy as an undefined technique which is applied to unspecified problems with non-predictable outcome. For this technique we recommend rigorous training.'

Luborsky attempts to gain sympathy for his plea by saying that 'if medical doctors had followed such advice their entire science would not have developed'. This statement ill agrees with my reading of medical history: I cannot recall any reputable method of treatment in modern times, used as widely as psychotherapy, and advocated as enthusiastically, which has been in use for over fifty years without any serious attempt to prove its effectiveness. This fact in itself is an interesting social phenomenon on which much could be said.

In summary, it appears that Luborsky has misunderstood the logic of my analysis, and that he does not, in fact, disagree with my conclusions. His disagreement with deductions made from these conclusions may be based on genuine differences of opinion, or on the fact that the major premise of my argument was not explicitly stated. If the latter, then it may be assumed that we are now in agreement; if the former, then we disagree in the sense that I regard psychology as a scientific discipline, while Luborsky does not. On this issue the reader no doubt will form his own opinion.

(*Manuscript received* 5 *December* 1953)

CORRESPONDENCE

THE EDITOR
British Journal of Psychology

Dear Sir,

Recently, M. M. Lewis presented a critical evaluation of our monograph on *The Acquisition of Word Meanings*.[1] The present note is intended to correct misapprehensions which may have been occasioned by this criticism. For this purpose, it will be necessary to sketch briefly the aim of the study, the nature of the method employed, the rationale underlying the testing of children of different ages, and the sorts of analyses undertaken.

The problem which we posed for ourselves—a problem grossly neglected by psychologists, experimental and otherwise—was the determination of the nature and variety of semantic processes as a function of developmental level. To this end, a special methodology was devised and children ranging in age from 8 years 6 months to 13 years 5 months were tested.

The method involved the establishment of a miniature model of the sort of situation confronting anyone who might seek to arrive at the meaning of an unknown word appearing in a variety of verbal contexts. The technique utilized artificial words, with six sentence contexts for each word; the subject's task was to give meaning to the artificial word on the basis of the clues offered by the various contexts. More precisely, the subject had to give the meaning of the artificial word, first in terms of the initial sentence alone, then in terms of the first two sentences, then in terms of the first three, and so on. With each addition of a new verbal context, the subject could maintain or change his solution; in either case, he was obliged to give his reasons for so doing, and was also required to show how the solution fit all of the contexts previously presented.

In the main, the series were so constructed that a single word, taken literally, could fit all six contexts. For several of the series, however, the English language did not provide a single word which could be taken literally for all six sentences. For example, the series containing the artificial word 'soldeve' consists of the following six sentences: (1) The dinner was good but the fruit we ate was soldeve. (2) When we were driving in the evening we did not feel safe because things on the road seemed to soldeve. (3) The older you get the sooner you will begin to soldeve. (4) People like a blossoming plant better than one that is soldeve. (5) Putting the dress on the sunny lawn made the colour soldeve. (6) Because the windshield was frozen things looked soldeve. An adequate solution for this series would be 'fade' or 'wither'. A solution in these 'unorthodox' cases had to be somewhat stretched if it were to fit all of the verbal contexts in a series. For such series, as for all of the other series, the correctness of a solution was operationally determined by the performance of educated subjects. In these 'unorthodox' series as in the other ones, there was a high degree of agreement among college students as to final solutions.

This point is stressed because some criticism has been directed against the entire test on the grounds that several of the series permit no adequate solution. Such criticism might be relevant if it were suggested that the analysis were directed toward such questions as the point within a series at which a correct guess is made or, indeed, whether a correct guess is made at all for any particular context. These questions were not involved in the study. In so far as correctness entered at all—and it must be emphasized that correctness of solution was merely a secondary aspect of the analysis—it was in terms of the degree of approximation toward a conventionalized solution as defined by a consensus of educated adult subjects. In any case, the criticism, which seems to assume that the main concern of the experiment was with correct guesses, misses the major aspects of the investigation.

---

[1] *The Acquisition of Word Meanings: a Developmental Study.* Reviewed in *Brit. J. Psychol.* XLIV, p. 269.

The experiment was chiefly directed at determining the processes involved in arriving at a solution. It was clearly evident that the same sorts of processes were utilized by children of a particular developmental stage in arriving at a solution for the 'unorthodox' series as were utilized in achieving solutions for the more orthodox series.

Now, with regard to the rationale underlying the employment of subjects differing in chronological age. Since many workers in the field of educational psychology are interested in establishing norms for specific chronological ages, it is important to stress that age was utilized in our experiment not in an absolute chronological sense, but rather as a relative index of development. In this respect, our approach bears a very close relationship to the manner in which Piaget utilizes age as an index of development. In other words, we used age with the assumption that the processes utilized by younger children to cope with cognitive problems are as a rule developmentally less mature than the processes utilized by older children. Whether certain processes are brought into play at a particular age will depend of course on the nature of the problem, the amount of experience of the subjects with the task, etc. Assuming these to be constant for all of the subjects, one can determine then what processes do emerge at various developmental levels to cope with the same experimental situation.

With regard to the analysis, we were predominantly interested in the nature and variety of the processes involved in giving meaning to words—not in the meanings *per se*, i.e. we sought to determine how children of different developmental stages handled the artificial word for each and every one of the six contexts comprising a series.

To give some idea of the findings of the study, only one aspect of one of the areas treated in the study may be mentioned, viz. that concerned with the nature and variety of the processes of signification. In the analysis, some fifteen different processes of signification were distinguished and ordered in a genetic progression. Statistical treatment revealed clear-cut evidence of a changing relationship, as a function of developmental level, between word-meaning and contextual involvement: for the youngest children, words and their sentence contexts were typically undifferentiated as to meaning—the word meant part or whole of the context in which it was embedded; with genetic progression, the word became increasingly detached from any particular context and finally reached the stage of a truly generalized concept. This and other findings are presented in considerable detail in the monograph.

In conclusion, we are convinced that the study does provide considerable information concerning the nature and variety of semantic processes in children, and also affords grounds for ordering these processes in a rational genetic sequence. The sort of approach we have utilized and the results obtained have been and are being fruitfully exploited in a series of experimental studies on the semantic activity of educated adults and uneducated adults, schizophrenics and feeble-minded children (Kaplan, 1950; Baker, 1953) The findings have also afforded an experimentally grounded basis for the interpretation of psychological activity involved in semantic changes in the history of the living language (Werner, 1954).

Yours faithfully,

HEINZ WERNER and EDITH KAPLAN

Clark University
Worcester
Massachusetts

REFERENCES

KAPLAN, B. (1950). A comparative study of acquisition of meanings in low-educated and high-educated adults. Unpublished M.A. thesis.

BAKER, R. W. (1953). The acquisition of verbal concepts in schizophrenia: a developmental approach to the study of disturbed language behavior. Unpublished Ph.D. dissertation.

WERNER, H. (1954). A change of meaning: a study of semantic processes through the experimental method. *J. Gen. Psychol.*

THE EDITOR
British Journal of Psychology

Dear Sir,

I welcome the opportunity of replying to this letter, as I was unable, in the space suggested by the Editor for my review, to substantiate my criticism as fully as I had wished.

I must repeat, what is made clear in my review, that I regard the general intention and the general planning of the investigation as fully satisfactory. There is work here to be done, since there are great gaps to be filled in our knowledge of the semantic development of children. It is all the more to be regretted that the authors should have given so much time to an investigation, the foundations of which are so insecure as to render any conclusions suspect.

Everything in this investigation depends, it is clear, upon the nature of the test-situations with which the children are presented. There are certain conditions which need to be satisfied; these are stated by the authors themselves in terms which seem to me unexceptionable. They say:

> The present study endeavors to investigate experimentally the processes underlying the acquisition of word meaning through verbal contexts. For this purpose the 'Word-Context Test' was designed. This test employs artificial words embedded in sentences. The subject going from one context to another is expected to arrive finally at the meaning of the word. Each of these artificial words signifies either an object or an action varying in degrees of concreteness. There are twelve series of six sentences each. The sentences in each series are, in general, ordered in such a way that, as a child moves from one sentence to the next, the clues increase in definiteness. (p. 1.)

In describing the administration of the tests, they say:

> The child was informed that he would be presented with twelve words which he never heard before; that these words were used in a little town out West and were not spoken anywhere else. It was made clear to him that each word had only one meaning throughout the six sentences. (p. 7.)

It is most unfortunate—and the justification of my criticism—that these conditions were not sufficiently satisfied in the devising of the test-situations. The intention was to present each unfamiliar word embedded in six English sentences; obviously it is essential, first, that there should be, as the authors say, 'one meaning throughout the six sentences'; or, as I would put it, at least semantic continuity; and, secondly, that since these are

English sentences presented to English-speaking children, there should at least be the possibility of the normal English range in the grammatical functions of the unfamiliar word from sentence to sentence. If, as is here the case, these conditions are not satisfied, there is the greatest likelihood that the children's responses will be determined not only by the manner in which they respond to unfamiliar words in meaningful contexts, but also by an immeasurable factor: the semantic and grammatical incompatibilities of the unfamiliar word from sentence to sentence. And the more literate the children are, the more likely are they to be confused by these incompatibilities.

To illustrate these points, I would draw attention to the *soldeve* example, cited in my review, and given in the letter above. The authors now call this an 'unorthodox example' but—as far as I can tell—there is no reference to its unorthodoxy in the monograph. I would also instance the following examples:

Hudray

1. If you eat well and sleep well you will hudray.
2. Mrs Smith wanted to hudray her family.
3. Jane had to hudray the cloth so that the dress would fit Mary.
4. You hudray what you know by reading and studying.
5. To hudray the number of children in the class there must be enough chairs.
6. You must have enough space in the bookcase to hudray your library.

Contavish

1. You can't fill anything with a contavish.
2. The more you take out of a contavish the larger it gets.
3. Before the house is finished the walls must have contavishes.
4. You can't feel or touch a contavish.
5. A bottle has only one contavish.
6. John fell into a contavish in the road.

My contention is that with test-situations as badly devised as these, no analysis of the results, however detailed, can provide trustworthy information about semantic processes in children.

Yours faithfully,

M. M. Lewis

Institute of Education
The University
Nottingham

PUBLICATIONS RECENTLY RECEIVED

Social Psychology. By S. A. Asch. New York: Prentice-Hall Inc. 1952. Pp. xvi + 646.

Social Psychology. By Theodore M. Newcomb. London: Tavistock Publications. 1952. Pp. xi + 690. 30*s.* net.

Social Psychology. By W. J. H. Sprott. London: Methuen. 1952. Pp. xiv + 268. 20*s.* net.

It is a commentary on the state of the subject that a reviewer can have before him three books with identical titles but remarkably different content. If, which is nowadays almost unthinkable, three books entitled simply 'physics' or 'chemistry' were produced, one could be reasonably certain that they would all contain approximately the same material; or that if one of them differed markedly from the rest, it would either be bad or else a work of Einsteinian genius. But the three books under consideration are all, in their particular ways, good books. There is certainly something to be learned from each, but although they have their areas of overlap, they all teach something different. Social psychology, in fact, has not yet established its basic premises. Its controversies do not arise from indecision as to the next step to be taken from an established base (as was the case, for instance, in the controversy over the wave and the particle explanations of light), but from a fundamental indecision as to the nature of the discipline, the appropriate methodologies, and the whole conceptual scheme by which relationships of man to his environment may be approached. Writers who do not understand the, as yet, rudimentary nature of our social-psychological concepts and techniques, produce naïvely all-conclusive theories about human nature and society. But Profs. Sprott, Newcomb and Asch are well aware that we are only at the beginning of things, and it is this realistic humility, coupled with a sensitive appreciation of the work already carried out, which makes their writing of value. It is perhaps a backhanded compliment to say that their books will help subsequent social psychologists to render them out of date before many years have passed.

Prof. Sprott's volume is the slightest of the three, but for this reason may prove to be the most satisfactory introduction. Moreover, it has the merit of being written by a man who has no psychological axe to grind. His erudition as a philosopher and a sociologist has given Prof. Sprott an admirable detachment in criticizing the theories of others. There are, it is true, certain gaps—for example, little is said of social role, which is perhaps the current key concept of social psychology—but on the whole the field is covered with lucid impartiality. Profs. Newcomb and Asch in their far more massive tomes (each has over 600 large pages, a deterrent almost as formidable as the cost) also cover the field of social psychology with great amplitude, but they do so to demonstrate a particular approach: Prof. Asch's is a gestalt approach, while Prof. Newcomb's combines the insights of the neo-Freudian with those of Lewinian field theory, though it must be admitted that these meagre descriptions do scant justice to attitudes towards human affairs which are elaborated with such scholarship and penetration. Of the two, Prof. Asch's is rather the more coherent, but he has perhaps taken the easier path. But both works will be indispensable to the student whose interests, it may be, have been stirred by Prof. Sprott's more easily digested work.

But while praising the work of men so able, and while recognizing how much they have done to systematize (albeit in somewhat varied systems) our knowledge of human motives and behaviour, it must be admitted that there are many central problems which remain untouched. The main one is, perhaps, that being human we can never attain real detachment and must always view man's actions through a screen of values and preconceptions for which no scientific assessment is possible, since they are outside the range of science. The intransigeance of this difficulty is one which becomes very clear if one compares work carried out in countries in which prevail differing attitudes towards, say, the meaning of responsibility, love, sex and duty. It is, for example, perhaps within the American ethos to *believe* more than we do in the value of the social sciences, while this in turn may spring from a sanguine belief in human perfectibility, which is typical of a new and vigorous civilization. For the psychologist whose work lies mainly in the laboratory, these difficulties are less intrusive, but so soon as we touch upon social behaviour, the problems of our work reflect in part our own problems, but often in so subtle a fashion that we have the illusion of complete objectivity.

But the works under review go a long way towards establishing a base from which more intractable difficulties may be assaulted.

ADAM CURLE

Field Theory in Social Science. By KURT LEWIN. Edited by Dorwin Cartwright. London: Tavistock Publications Ltd. 1952. Pp. xx+346. 21*s.* net.

This posthumous collection of theoretical papers written by Lewin during the last ten years of his life is a companion volume to *Resolving Social Conflicts*, in which his chief concern was with practical problems in society. The papers are selected and edited by the present Director of Lewin's Research Center for Group Dynamics, who assures us from that vantage point that 'they add up to a remarkably comprehensive statement of his major contributions'. The book begins with three essays in the philosophy of science. These three chapters, together with a technical appendix, are a declaration of Lewin's theoretical faith and a detailed—often exhilarating— exposition of its articles. The papers forming the other three-quarters of the book show how these guiding principles apply to a wide variety of specific problems in social psychology and neighbouring fields. The last chapter, reprinted from Carmichael's *Manual of Child Psychology*, summarizes the main theoretical and empirical findings of research carried out under Lewin's supervision.

Had Lewin written a book with field theory in social science as his theme it would certainly have been very much like this assembly of papers, but better. He would undoubtedly have developed his argument with a continuity, fullness and coherence which one has really no right to expect in a collection of separately conceived discussions, but the 'out-group' reader of this one may become a little daunted as he moves on and finds it repetitious, rather disjointed and occasionally over-compressed. However that is after all a minor matter of organization—unavoidable probably in this instance but odd, nevertheless, in a publication presenting the work of someone who so persuasively emphasized how important for a whole is the interdependence of its parts! In style and substance the individual papers are authentic and essential Lewin. As such they display his distinctive characteristics in all their strength and weakness, and their value is enhanced by the fact that all were written after the publication of the three books which have made his name in the English-speaking world. With so much first-hand and up-to-date material at his disposal the reader will feel better equipped to appreciate Lewin's theoretical contributions to social science, and any chronic sufferer from 'fact-happy' empiricism could hardly find a better antidote than these wide-ranging writings of a dedicated theorist with a flair for tackling practical problems.

A. J. FLOOK

Instincts of the Herd in Peace and War. By WILFRED TROTTER. Edited by R. W. CHAPMAN, with a foreword by F. R. M. WALSHE. Oxford University Press. 1953. Pp. xviii+220. 25*s.* net.

It shows the remarkable vitality of Trotter's work that this library edition of his book (prepared by R. W. Chapman, with minor corrections from Trotter's manuscripts and typescripts, and with a new and more comprehensive index) should be published while an ordinary edition is still in print. The book forms the best kind of memorial to the author, and Dr Walshe's Foreword, giving some account of Trotter's personality, shows him to have been one of the old civilized men of science, with a mind capable of interesting educated people. Dr Walshe regrets that although sympathetic towards religious experience and alive to its importance, Trotter viewed it detachedly and without having a religious faith himself.

Trotter's technical formulation of his insights is not acceptable to contemporary psychologists and could undoubtedly be improved. But he *has* insights, of which there are none too many in psychology, and he sees the relevance of his insights to the practical affair of living a civilized life. The astringent detachment of his writing is well illustrated in his remark: 'To suppose that when one has admitted the liability to prejudice one can free oneself from it by a direct voluntary effort is a common belief and an entirely fallacious one. Such a task is far beyond the powers of the most fully instructed mind, and is not likely to be undertaken except by those who have least chance of success'. Reading to-day, we have to make great allowances for Trotter's own wartime prejudices and for the pseudo-biological form in which they were expressed; he achieves much more detachment in the interesting and prophetic 'Postscript of 1919'. There are also obvious dangers in his conception of socialized gregariousness on the ant or bee pattern as an ideal for human communities. But to make his mistakes and dangerous tendencies an excuse for avoiding the stimulating challenge of the book would be to betray oneself as one of the 'stable-minded' to whom Trotter attributes responsibility for our worst social mishaps, including war.

D. W. H.

Solitude and Privacy. A Study of Social Isolation, its Causes and Therapy. By PAUL HALMOS. With a foreword by T. H. MARSHALL. London: Routledge and Kegan Paul Ltd. 1952. Pp. xvii + 181. 21*s.* net.

The title of this book is deceptively modest. Dr Halmos does not merely discuss the phenomenon of 'social isolation' in terms of some standard psychological theory. He is concerned rather to put forward a new fundamental theoretical principle, namely that psychology needs to recognize the existence in man of a definite need for 'social participation' or fellowship, not merely as a secondary manifestation of the sex-drive, as in Freudian theory, but as a primary drive in its own right, rooted in man's biology as being an evolutionary factor with considerable survival-value. The recognition of such a '*principium sociale*' operating alongside the '*principium sexuale*' (the rather irritating jargon is one of this book's less fortunate features) would clear up, Dr Halmos claims, a number of difficulties in the Freudian and neo-Freudian theories, and he points out that something very like it has been postulated in different ways by such diverse psychologists as Rank, Fromm, Sullivan, Adler and Jung. In terms of it an interesting diagnosis of some of our present discontents becomes possible, for we live in a 'desocialized culture', where opportunities for social participation such as were afforded in past centuries by choral dancing and communal drinking have declined (statistics are provided to prove this) and an 'ideology of privacy and reserve' has grown up to confirm people in their separateness (passages are quoted from many 'representative' modern writers to illustrate this). While agreeing that in particular cases inability to 'participate' in social life will be a result of neurosis as ordinarily understood, Dr Halmos follows American psychologists of the type of Mead in holding that in the final analysis 'socio-cultural determinants precede and control individual-genetic development', so that desocialized culture must be regarded as one of the primary causes of individual neurotic development itself. He therefore concludes that individual psychotherapy will have only very limited success unless it takes place within the context of a much wider 'community therapy', by which he means not social reform of the ordinary kind, but the deliberate attempt to create fellowship-groups into which people can enter and find satisfaction.

There is little doubt that Dr Halmos is getting at something extremely important, but his argument begs some large questions, even for those readers who would be willing to agree that he has made his case for some sort of need for relationship in human beings over and above secondary manifestations of sexuality. The principal question is one of phenomenology and logic. He uses the term 'friendship', or his own term 'biosocial participation', indiscriminately to cover all group-relationships except mechanical regimentation, and seems unaware that there are differences between groups so great as to make nonsense of this. Can the loss of personality in herd-relationship, such as comes from the tribal dance or alcoholic intoxication, really be said to satisfy the same need as friendship in which personality is enhanced? Or, to put the point in another way, does not the demand for privacy itself express a fundamental need of man, the need to escape from the 'dear octopus' of the herd in order to become fully human? Is not privacy indeed the precondition of real friendship? It is to be hoped that Dr Halmos's book will stimulate further reflexion and research on the important question of the *differentia* of relationships between persons, without which community-therapy is liable to be a groping in the dark. JOHN WREN-LEWIS

Sigmund Freud. Life and Work. Vol. I. *The Young Freud, 1856–1900.* By ERNEST JONES. London: The Hogarth Press. 1953. Pp. xiv + 454, with 10 illustrations. 27*s.* 6*d.* net.

This, the first, volume of a biography which will doubtless rank as classic in the eyes of posterity conveys the impression of being above all a labour of love on the part of the author; but it is conceived in no spirit of hero-worship. It is conceived rather in a spirit of admiration tempered by the most realistic appraisement of the emotional factors influencing Freud's intellectual achievements. In view of the outstanding nature of these achievements, the appearance of the present record is particularly timely; for, since Ernest Jones is the last remaining of the inner circle of Freud's collaborators, there is no other who could ever speak with the authority which he commands. What has influenced him most in the preparation of the biography is 'the hope of elucidating something of the processes in Freud's mind, and the experiences in his life, that culminated in his discoveries'. In a sense, therefore, the biography may be regarded as representing a psychoanalytical study of the origin of psychoanalysis. One of the most important sources drawn upon for this study is the remarkable collection of intimate letters written by Freud to his friend Wilhelm Fliess between 1887 and 1902—a period covering the most original phase in Freud's thought; and, as Jones convincingly shows, there was an interreaction between Freud's intellectual productivity at this time, the psychoneurosis from which he is now known to have suffered, and his emotional relationship with Fliess. The picture of Freud which emerges, so far from conforming to the popular picture of the detached observer calmly elaborating theories about human nature, is that of a sensitive soul heroically struggling to solve his own emotional problems, and in so doing to achieve scientific understanding of the mainsprings of human motivation. It was in the course of this struggle that Freud was inspired to embark upon the arduous task of self-analysis, which resulted not only in his achieving the remarkable personal adjustment which characterized his later years, but also in the conceptualization of his insights in the psychoanalytical theory which has made his name famous. The enormous labour undertaken by Jones in exploring the sources of the intellectual and emotional influences which bore upon Freud, and in relating these to the theoretical system of psychoanalysis which emerged under the influence of Freud's genius must be regarded as amply rewarded; and the reader of the first volume will look forward to subsequent volumes with eager anticipation.

W. RONALD D. FAIRBAIRN

On Aphasia. By S. FREUD. Translated by E. STENGEL. London: Imago Publishing Co. Ltd. 1953. Pp. xvi + 105. 12*s.* 6*d.* net.

Dr Stengel has performed a most valuable service in translating this fascinating little early work of Freud. Long since rare in German, and never before translated into English, it now takes its place as an important link in the history alike of psychoanalysis, and of the study of disordered language function. That Freud drew some inspiration from Hughlings Jackson must have suggested itself to everybody who has seriously puzzled over the idea of 'regression' on the one hand, and that of 'dissolution of function' on the other. In this book Freud explicitly bases his constructive views about aphasia upon those of Jackson. In doing so he was a pioneer—preceded, it seems, only by Arnold Pick. The interest of these two appears to have found no response, until Henry Head republished Jackson's papers in 1915. It is an intriguing fact that Head, although he mentions Freud as the inventor of the term *agnosia*, gives no evidence that he had ever read Freud's book, yet this, alone of previous works—with the exception of Pick's papers—propounded the need for a functional and psychological, as opposed to an anatomical, interpretation of the phenomena of aphasia along lines directly suggested by Jackson, and so brilliantly developed by Head himself.

That these ideas, derived from Jackson in connexion with aphasia in particular, and the dissolution of nervous functions in general, exerted a profound influence upon the later development of Freud's thought is evident. Freud characterized his psychoanalytic theories as 'superstructure which will have to be set on its organic foundation at some time or other'. Jackson did not hesitate to extend formulations rooted in neurological fact into the field of psychological observation. The lines of thought which they respectively developed are perhaps among the most significant that offer hope for the future. It is pleasing to learn from this book that these two men were not mere contemporaries.

R. C. O.

Ego Psychology and the Psychoses. By PAUL FEDERN. Edited and with an Introduction by EDOARDO WEISS. London: Imago Publishing Co. Ltd. 1953. Pp. 375. 30*s.* net.

Paul Federn trained under Freud and was closely associated with him from 1903 onwards until he left Vienna and settled in New York in 1938. He was one of the original Freudian circle of eight. He died in 1950 at the age of 79.

The book is divided into three parts: (1) on ego psychology; (2) on the treatment of psychoses; (3) two papers on narcissism. There is an excellent introduction by Weiss, a bibliography of Federn's publications and an index. Most of the chapters are reprinted from Federn's published writings, but a few are from manuscript or compiled from his papers. They date from 1926 to 1949. As one would expect, they are varied in length and a great deal of repetition was inevitable. It would have been an advantage if the dates of the various publications had been given in the list of contents. If these dates are looked up in the footnotes at the beginnings of the chapters, or in the list of Federn's works, it is found that the chapters in each of the three sections of the book are in order of publication, except those now published for the first time.

Part two of the book, on the treatment of the psychoses, is the most readable and easy to understand, and will be of special interest to psychiatrists and psychotherapists who are dealing with the psychoses. Federn had much experience with the problems of the treatment of the psychoses. In 1906 he started experimenting on the therapy of what were regarded as hopeless cases, knowing that he could not do them any harm, and gradually worked out therapeutic techniques for dealing with them, especially for schizophrenia. He shows that Freud was right in saying that psychotic states cannot be treated by psychoanalysis only in the sense of the classical type of analysis, which was formed to suit neurotics, but that it can be modified in such a way that with the insight given by ego psychology it can also be used therapeutically for psychotics. Many people are now working on the problems of the psychoanalytic therapy of psychotics, and therefore the book will excite considerable interest and be of much value.

The introduction by Weiss is very helpful, and it is rationally correct that the book should start with a theoretical part which explains the ego psychology with which Federn works. In practice, however, the presentation of this part, the ego psychology and its implications, is very abstract and will be difficult for the reader, and it would have been better to start with the second part of the book, which is more concrete and is presented in such a way that the reader can see why the special theories about psychotic cases should have been formulated. It is more natural to step from problem to theory and then to practice, than directly from theory to practice. In any case the book inevitably suffers from the mode of its construction out of a series of papers written separately for various occasions and purposes, and it would have been very much more satisfactory if it could have been wholly rewritten. Then the various steps of theory could have been integrated with concrete material and examples showing how psychoses, latent and partial psychoses and pre-psychotic states were handled in the therapeutic situation. If this had been done, the theories relating to ego psychology could have been inter-related with other branches of psychoanalytic theory, such as the work of Melanie Klein, which Federn considers of special importance but never discusses.

Federn points out that every psychosis is a mental disease of the ego itself, while in neurosis only some functions of the ego are impaired. In his handling of the psychoses his practices are not difficult to grasp, and are closely related to methods which have occurred to many of us independently in our approaches to partial and latent psychotics and to the psychotic aspects of some disturbances which would be classed as neurotic in conventional diagnosis. Federn appreciates this, but claims that his ego psychology is a direct derivative from the truly Freudian plan, and that it will enable us to treat these problems with clear scientific insight and knowledge of what is being done to the patient, rather than by methods which though often more successful than classical psychoanalysis for certain patients, are essentially unscientific or indirectly derived from the Freudian line.

His formulations of the processes involved in psychotic disturbances are brilliant and fascinating, and they show him to be a man of exceptional insight. The applications of the ego psychology itself, however, are not always quite convincing. This is because he usually leaves out the analysis of the deeper unconscious mechanisms at work in his efforts to concentrate attention on the all-important ego-aspects which are neglected by most other writers.

In general he explains that schizophrenia and schizoid states arise from breakdown and inadequate libido cathexis of the inner boundaries of the ego, so that it is invaded by deeply repressed material and impulses and also loses its firmness and coherence. Depressive states are in general due to the cathexis of the ego boundaries with mortido or the death instinct. In the psychoses there is a tendency for ego boundaries to be withdrawn and to regress to earlier and more infantile forms. In this way in paranoid states thoughts are sometimes excluded from the ego boundary and assume a false certainty, In schizophrenia external reality is confused with falsified ideas about reality, while the tendency of the inner boundaries of the ego to dissolve gives rise to feelings of depersonalization which are intensified by the recathexis of earlier and more infantile ego boundaries, so that there a multiplicity of ego forms cathected at the same time. In paranoia there is a falsification of thought sequences and the individual loses his capacity for logic and self-criticism. The character of planning explains the aggressive and systematic qualities of paranoid thinking. The manic patient continually abandons his pleasure-giving objects and ideas for new ones, as if they were sources of unpleasure.

In all these ideas (which are explained much more fully in the book and cannot be treated with justice in a few sentences) fascinating though they are, the ego psychology tends to leave us with momentous questions in our minds. Why do all these tremendously powerful ego-changes take place at all? What makes the ego withdraw its boundaries and regress to earlier states? Why are the ego boundaries sometimes invested with mortido? Why should the inner firmness of the ego dissolve and what is the nature of the powerful impulses and mental contents which then invade it? Why does the manic continually escape his love objects as if they displeased him? Why does the paranoiac substitute false steps of such vast moment in his thinking, and how can the quality of planning explain its savage but rationalized vindictiveness? If we turn to Melanie Klein we shall see to some extent how these and many other difficult questions might be answered.

There are many points in the book which might be taken up for discussion by a thoughtful reader, but no space can be found for them all here. For instance, it seems that the essential method of ego-therapy of the psychoses consists in helping the patient to come to insightful terms with his problems, and to re-repress the invading unconscious impulses and material, instead of lifting the barriers of repression as in treating the neuroses. It is therefore perhaps more of the nature of rehabilitation than of radical cure, and possibly this is inevitable and may be one reason why group therapy is specially helpful to psychotics. A difficulty occurs early and runs throughout the whole book, in that Federn appears to think that the psychoses are partly determined in an essentially biological and hereditary way, but the problem of the interaction of heredity and environment in the causation of the psychoses is never fully discussed.

In spite of various criticisms which are not difficult to make, the book will excite much interest and deserves re-reading many times. It will introduce many readers to Federn's life-work which will provide a most valuable step forward in the understanding and treatment of the psychoses, together with partial and latent psychotic states and those essentially psychotic aspects of the neuroses which cause every analytic therapist particular difficulty when they are encountered. R. W. PICKFORD

Brett's History of Psychology. Edited and arranged by R. S. Peters. London: George Allen and Unwin Ltd.; New York: The Macmillan Company. Abridged one-volume edition 1953. Pp. 742. 42*s*.

The re-issue of Brett's *History of Psychology* in an abridged (but still hefty) one volume edition raises some interesting questions about the different kinds of reader to whom the longer and the shorter versions may appeal.

Brett's original history, published in three volumes between 1912 and 1921, was a masterpiece of erudition, providing a survey of psychological theorizing from classical Greece to the end of the nineteenth century. He wrote as a psychologizing philosopher and often a very good one. The need to summarize made him sometimes a little sleep inducing. But he could be caught talking to himself in a way that was mature, humorous, insightful and often pithy. He rescued odd gems such as the probable first English use, in 1665, of the word 'anthropologie' in a book which divided the subject into psychology and anatomy. The writer announced that 'of the former we shall in a distracted rehearsal deliver our collections'. Brett clearly had a sneaking sympathy for his anonymous forerunner and one can understand why his students found him a stimulating and charming teacher.

Of his generation, and averse from talking down, Brett expected of his readers enough general background to pick up often cryptic allusions, without which equipment the significance of what he said was not always obvious. How much social and political history, for example, lies behind the remark that Descartes 'lived in close enough contact with beliefs about human souls taking up their abode in animal bodies to feel the immense advantage of a more scientific view of the matter'. There are stretches of the original Brett where the present reviewer is uncertain whether to criticize or to confess her own ignorance. Sometimes, however, the mid-twentieth century reader would wish to take Brett to task for not giving more of such history *en clair*, and showing its relevance to the cultural tradition in which modern psychology grew—and for that matter is still developing. Brett's medieval thinkers, for example, need the setting provided by an H. O. Taylor, a Huizinga or an Eileen Power. His Machiavelli and Pomponazzi droop a little unless the reader is prepared to meet them with some understanding of Italian Renaissance politics. Brett might be criticized, on the other hand, for having brought in too many minor deities, so that the wood is rather hard to see for the trees and there are details that could have been cut in the interests of someone wanting a broad outline.

Finally, and not irrelevantly, the original version, valuable reference book that it was, suffered because Brett was erratic in supplying immediate details of date and reference. Many of the latter he gave in notes at the end but his volumes were not handy for rapid consultation.

What have been the effects of Dr Peters' editing? First, we have now got one book instead of three and it is available instead of being out of print. Secondly, some of the sensible abridgement has helped to clarify the picture. Thirdly, by the cuts, new sections in clearly different type, new chapter headings and a final chapter on the twentieth century Dr Peters has given the book a rather different focus. The history as he produces it shows, more clearly than the original, changing emphases in the criteria of acceptable explanation and scientific method. To readers oppressed by Brett's philosophizing and detail, and barely convinced that the past of psychology has any relevance to understanding the present, these changes will be advantageous. Others may feel that the price of change of emphasis was too great a one to pay for the slant that might have been better achieved by the writing of a new book.

Dr Peters had a very difficult job. He shows he is well aware of the divergence between our own and Brett's approach when he comments that 'to many readers the revised edition will, perhaps, appear more like the Festival of Britain Skylon sticking out of the Crystal Palace than the winning design of Coventry Cathedral'. The present reviewer found the 1951 Skylon attractive and Dr Peters' counterpart is often useful and effective. Perhaps in a structure of his own making it could be a little more graceful and his plea for recognizing the importance of adequate hypotheses might then be supported by his own suggestions of good ones to try out.

Dr Peters' cuts and section headings, justified as they often are, will have made it even more necessary for new readers to be armed with ordinary history. He has not rectified Brett's omissions of useful dates. His substitution, for Brett's notes, of a bibliography, in itself helpful, means that the reader who wants the original sources may still need to go to the older edition. The book could do with an index of subjects as well as names.

In sum: the new Brett is a different Brett, retaining a great deal (over 600 pages) of the original material, including many of Brett's most delectable irreverencies. It may attract an audience to which its predecessor will have been closed because appearing out of date and alien. It would be a pity if such readers, once tempted, did not then look up the older man, and with him do a little more exploring. It will be a great pity if Dr Peters confines himself to a piece of onerous editing and does not write the book of his own of which he is clearly capable. JOAN WYNN REEVES

Cerebral Mechanisms in Behavior. Edited by LLOYD A. JEFFRESS. London: Chapman and Hall Ltd; N.Y.: John Wiley and Sons, Inc. 1951. Pp. xi+311. 52*s.* net.

The Hixon Symposium, held at the California Institute of Technology in 1948, was distinguished by the eminence of its members, each in some particular approach to the problems of cerebral functioning. The six main papers in the published version have been revised by their authors; the discussion, which occupies nearly as much space, has been edited as little as possible—a policy which, although fairly certainly the best course, adds the difficulty of following a shifting debate to that inherent in the complex subject-matter.

The papers fall into two groups. The first, including papers by Von Neumann and McCulloch (one by Lorente de Nó could unfortunately not be published), argues deductively towards behaviour from hypothetical neural mechanisms, which are themselves derived from the elements of computing machines. The result, as several contributors point out, falls short of behaviour as empirically observed; and the theories proposed, though provocative, and perhaps capable of some degree of physiological confirmation, do not appear as yet to lead to hypotheses which might be verified in behaviour.

The procedure of the contributors in the second section is rather to argue back from observed behaviour to the mechanisms which might underlie it. Lashley considers the problem of the temporal integration of behaviour sequences in the light of current knowledge of reverberatory circuits; Klüver discusses functional differences between the occipital and temporal lobes as revealed by behaviour changes following extirpation; Köhler reports electrophysiological observations made in the course of his studies of figural after-effects; Halstead deals with the relationship between brain-functions and intelligence.

This presentation of a sample of the most advanced contemporary thinking in a difficult field leaves one with the feeling that we have not as yet sufficient understanding of either behaviour or cerebral mechanisms for one to throw any great amount of light on the other. This conclusion should not, however, be taken as a reflexion on the present volume, which clearly merits the attention of anyone interested in the development of physiological psychology. N. E. L.

(The Editor regrets the delay, for which the reviewer was not responsible, in the appearance of this notice.)

An Introduction to Experimental Psychology. By C. W. VALENTINE. Fifth edition. London: University Tutorial Press Ltd. 1953. Pp. viii+272. 8*s.* 6*d.* net.

The appearance of the fifth edition of this book, which was first published forty years ago, is testimony to its soundness and its acceptability to successive generations of students and instructors. It is organized in two sections—Part I now containing details of no fewer than fifty-seven different experiments (compared with twenty-seven in the original edition), Part II discussing the problems studied and the results obtained. The range is very comprehensive but with a bias towards educational applications. There is nothing, for instance, on emotion or on sensation or the psychophysical methods. There is nothing of brass-instrument psychology about it. All the necessary materials for the experiments are provided in the text or can be easily prepared. It is, avowedly, an elementary book. Only the simplest statistical treatment of results is called for. The experiments are simple but demand thorough and careful work. Any student who has worked through half of them, read the discussions and followed up the recommendations for further reading in a bibliography which has been brought right up to date, will have a really solid foundation for more advanced work and, almost certainly, a better understanding than one who has plodded through a textbook without doing any practical work. Prof. Valentine recommends, of course, that his book should be used in conjunction with a textbook of general psychology. New experiments in the field of learning and perception and two corresponding chapters in Part II have been added in the new edition and the book now appears in a more modern format. The manual may be warmly recommended to those who are not already familiar with it and will continue to have a wide field of usefulness.

A. H. D. T.

Uses and Abuses of Psychology. By H. J. EYSENCK. London: Penguin Books. 1953. Pp. 318. 2*s*. 6*d*.

This book comprises a non-technical account of the impact of applied psychology on human affairs. The author is vigorous in his defence of the subject as 'a science in its early formative period', candid in his readiness to say 'we don't know', and insistent that the right to criticize scientific work 'must not be assumed lightly' by the layman. Dr Eysenck is highly critical of those areas of psychology where research findings do not seem to justify present applications, though consistently constructive in his argument that indifferent work must, and ultimately will, be replaced by better. An interesting feature is the confidence shown towards the future development of a more rigorously scientific social psychology.

Psychologist readers may wish to take issue with Dr Eysenck on a number of matters, including quite fundamental ones. Those who regard science as concerned with a continuum of probability values may well dislike the author's frequent dichotomy as between 'no evidence' and 'adequate proof'. Although the book's emphasis upon the need for verification and falsification of alternative hypotheses in this field is to be welcomed, the treatment of psychoanalytic work seems unnecessarily severe. Nevertheless, few will wish to oppose reiteration of Freud's own sober advice against 'indiscriminate psycho-analysing'. Dr Eysenck is less than just to Freud whom he accuses of requiring that one 'believe in his system before one can criticize it', though Freud is doing little more than insist on adequate specialist training for psychoanalysts; Dr Eysenck's own standards of specialist training, for psychologists, are similarly exacting. Again, it is not always clear to the layman when the book is stating accepted as opposed to controversial conclusions. Many psychologists might agree, for instance, that the usefulness of psychotherapy has not as yet been scientifically proved, and that more research is needed. Without further evidence few would be willing to endorse certain other views expressed, for instance the tendency to explain predisposition to neurosis by genetic factors rather than genetic and early learning components together. Among hopeful signs of development is the readiness of Dr Eysenck, as a leading factorial investigator, to identify himself so closely with the standpoint of an experimental investigator of learning—in this instance Mowrer. A second encouraging sign is the author's similar sympathy towards psychoanalysts who *have* been very much concerned to verify their hypotheses (for instance the contributors to *The Authoritarian Personality*, one of the ten books recommended for further reading). Despite this, Dr Eysenck remains highly critical of psychoanalysis in general, his main objection being on methodological grounds.

In the interesting account of intelligence theory and testing the author is in close accord with the standpoint of Thurstone, but also reports some illuminating researches by Furneaux with some quite fundamental implications in this field. A convincing account is given of the handicaps which confront the psychologist, but not the physical scientist, in investigating a problem. Again, the Royal Commission which seeks and accepts mere opinions, whether specialist or otherwise, instead of conducting a scientific investigation, is heavily criticized. Finally, in political matters, Dr Eysenck strongly supports scientific inquiry as an independent alternative to doctrinaire attitudes of both the political Right and the political Left. The claims of a scientific psychology to play a larger part in human affairs have seldom been more ably defended. It is fortunate that this defence, as a *Penguin* publication, will be so widely read by non-psychologists.

PETER McKELLAR

Psychomatic Approach to Gynecology and Obstetrics. By FRITZ WENGRAF. Oxford: Blackwell Scientific Publications; U.S.A.: Charles C. Thomas, Springfield, Ill. 1953. Pp. xx + 346. 50*s*. net.

This book is a monograph in The Bannerstone Division of American Lectures on Gynecology and Obstetrics, under the editorship of E. C. Hamblen. It is a valuable contribution to the subject, from psychoanalytic and psychosomatic points of view. It deals with the psychological and psychiatric aspects of all female generative functions, and is addressed primarily to physicians, gynaecologists and obstetricians. Apart from the value it has to them, it contains much material of general interest to all clinical psychologists. There is a bibliography of 516 works.

R. W. PICKFORD

Introduction to the Psychology of Music. By G. Révész. London: Longmans, Green and Co. 1953. Pp. xiv + 261. 30*s.* net.

In Révész's own words, '...my main objective has been to introduce readers to the psychology of music—and at the same time give a systematic outline of my own research in this domain and that of acoustics.'

The book is divided into three parts: Part one deals with the physical attributes of musical tones; our musical system in terms of the different temperaments, with full lists of their ratios; the physiology of hearing; the theories of auditory perception and the physiology of voice. Part two describes the psychology of isolated tones and intervals outside the musical context, by the ear; the character of the different keys and the relationship between tone and colour. Part three is devoted to musical ability in its various aspects, including chapters on blind musicians, on the relationship between musical and other types of abilities, and on the inheritance of musical ability. Added are a few chapters only loosely connected with the psychology of music: on creative work; musical ability in the deaf; on the pathology of musical perception; the origin of music, and the relationship between aesthetics and music. The bibliography is restricted to books quoted in the text.

Taken as a whole, this book is a collection of articles on a wide range of subjects, some of which are only remotely connected with the psychology of music. Most of them vigorously polemic, they were written at different times, for a different purpose, and they vary in range, value and depth. Only by discussing each on its merits, could one do justice to them. The double object of writing an introduction to the psychology of music and of giving a 'systematic outline' of research extending over forty years, resisted the integration of heterogeneous material into a whole that could serve the purpose of both.

While stating that 'Music is not concerned with single notes and their relations as expressed in isolated intervals and chords', Révész devotes about one-third of the volume to these subjects. At the same time, he relegates the 'more complex tonal forms' to 'the realm of general music theory, aesthetics, or the laws of musical form and style'. The perception of music as gestalt, one of the fundamental problems of the psychology of music, is given little more than one page.

An introduction to the psychology of music should be a source of general information. It would, by its very nature, require an impartial and balanced presentation of the various aspects of the subject, complemented by a well selected and up-to-date bibliography as a guide for further information. For an introduction to the psychology of music this is a disappointing book. But the different chapters, each in its own right, make interesting and stimulating reading.

MARTHA VIDOR

Intelligence Testing and the Comprehensive School. By B. Simon. London: Lawrence and Wishart Ltd. 1953. Pp. 112. 8*s.* 6*d.* net.

This is an intelligent and persuasive exposition of the familiar left-wing criticisms of intelligence tests and secondary school selection. It is worth the careful attention of psychologists, since many have in the past been too prone to believe that their tests do measure inborn and unalterable ability. But the author neglects to mention the more widely accepted current view that any hereditary potential depends for its realization on a suitable environment; that measured intelligence is therefore the product of genetic and environmental factors, and that comparisons of groups or individuals brought up in widely different environments (e.g. civilized against primitive) are illegitimate. Although this implies that the I.Q. is to some extent inconstant and affected by social class, intelligence tests do nevertheless possess useful reliability and validity as selection instruments; moreover, social class differences are themselves partly genetic. The irrefutable scientific evidence for these statements is ignored, since it would of course contradict the party line.

Mr Simon's complaint that the psychologist's concepts, such as intelligence, are conditioned by his cultural ideology, is also quite sound—in the light of contemporary field theory. Indeed there is much to be said for discarding the term 'intelligence' tests, especially at the 10–11 year examination, and substituting tests of 'educational aptitude'. But his account of the controversies over the definition of intelligence, and of the solution provided by factor analysis, is muddled and misleading. Far from the various definitions being 'mutually exclusive', it is because they—or at least the tests devised to measure them—overlap so much that it is possible to isolate g, and progressively to analyse out the verbal, spatial and other content factors and such formal factors as speed, facility with new-type items, etc., which often enter into our alleged tests of intelligence.

Much is made, too, of the circularity that is sometimes involved in arguments about the normal distribution of abilities; for the extreme socialist cannot stomach the idea that only a limited proportion of the population has sufficiently high capacity to be able to tackle advanced secondary school courses. While it may well be true that present methods of streaming in primary and secondary schools do not give the more backward as good a chance to make up as more individual and flexible classroom organization, there was plenty of evidence of ineradicable intellectual differences in the days before intelligence tests or streaming in state schools had ever been thought of. And the claim that in Russia and Poland all children can be brought on at a relatively equal rate and achieve a much more equal level of educational attainment is so dubious that it can presumably never have been put to the test.

Many of the author's criticisms of the present system of secondary school organization and selection are justifiable. Its bad effects on primary school teaching are deplored by thoughtful educationists and psychologists. They would agree that there is no perfect method of selection for grammar schools, especially at such an early age as 10, and that numerous pupils relegated to modern schools (particularly in areas with low grammar school provision) might with a better education surpass some of those admitted. But to ascribe all these ills to the wicked psychologist and his class-conditioned concept of intelligence is ridiculous. Mr Simon nowhere attacks the parents whose class prejudices in favour of a grammar school education are largely responsible for the examination becoming so highly competitive. And he ignores the extremely complex historical factors and the political considerations which underly the present differentiation of our various types of secondary schools. It is probable that few of the people concerned with educational selection or guidance have so little influence on educational policy as the psychologist. Nevertheless, in the present writer's view, the time is ripe for psychologists to get together and attempt to decide how far their instruments are doing harm, and how they could be better employed. This is one reason why a rather slight book, which might be regarded by many as mere political propaganda, has been accorded a fairly lengthy review. Mr Simon has taken the trouble to study the technicalities of his subject far more thoroughly than most propagandists (though it is a pity that he relied so heavily on a book written by the reviewer 15 years ago); and his distortions of fact are remarkably few. He has made a serious case, and deserves a serious answer.

P. E. V.

Secondary School Entrance Examinations. By A. F. WATTS, D. A. PIDGEON and A. YATES. National Foundation for Educational Research Publication, no. 6. London: Newnes Educational Publishing Co. Ltd. 1952. Pp. 77. 3*s.* 6*d.* net.

This is the second interim report from the National Foundation on the allocation of primary school leavers to courses of secondary education. Its content is accurately indicated by the sub-title 'a study of some of the factors influencing scores in objective tests with particular reference to coaching and practice'. The authors survey a number of recent investigations into the effects of coaching and practice and describe some experiments of their own designed to answer certain definite questions about their influence. There remain some inconsistencies between the findings of different investigators which are not cleared up, and there are some puzzling results which are difficult to interpret. It is hard to see, for example, why both three hours and nine hours of coaching should consistently be more effective than six hours.

It is confirmed that the objective type of test is more reliable and fairer than any other kind of examination at the 11+ level, but it has not perfect reliability and is not always absolutely just in comparing pupils from different schools who have had varying amounts of preparation for the intelligence test. Further, it does not tell us all we need to know about the examinees and especially about the borderline cases. The authors disapprove of universal coaching for the intelligence test which has been recommended in some quarters and see no need for it. They advise sufficient practice to swamp the effects of differences in preparation—not a great deal is necessary. They have also a number of sensible comments and recommendations on the general problem of secondary school allocation and indicate where further research is especially needed.

A. H. D. T.

Shame and Guilt. By GERHART PIERS and MILTON B. SINGER. U.S.A.: Charles C. Thomas. 1953. Pp. x+87. 23*s.* 6*d.* net.

This slim, elegantly-bound volume of 86 pages contains two essays, one by Dr Piers, the other by Dr Singer. They are concise and intelligible, and may well serve as models for the less restrained and more obscure of American writers on social psychology. The points made by each writer are quite straightforward. Dr Piers insists that shame and guilt represent different tensions, the former between the 'id' and the 'ego-ideal', the latter between the 'id' and the 'super-ego'. It may be that guilt masks shame, as when a person is prompted by his sense of inadequacy to do something of which the super-ego disapproves, and then feels guilty; it may be that the reverse happens, as when a person seeks *shameful* punishment. This may make it difficult to detect the primary tension at work in any given syndrome, but the distinction is, nevertheless, of vital importance for adequate understanding and for therapy. The guilt-ridden and the shame-ridden have different personality structures.

Dr Singer, in his somewhat less conclusive essay, examines, in the light of what Dr Piers has said, the familiar distinction between guilt-ridden cultures, and shame-ridden ones. He shows that the distinction is by no means as absolute as it has appeared to those who wish to argue that the 'super-ego' is an invention of Western cultures. Having shown this, he unfortunately does not go on to discuss the causes of the differential emphasis (rather than clear-cut difference) which he seems to admit. However, his arguments against those cultural anthropologists who want to make a more extreme differentiation between cultures than the evidence warrants will be of great interest to all social psychologists, interested in alternative methods of socialization.

W. J. H. S.

The Deprived and the Privileged. By B. M. SPINLEY. London: Routledge and Kegan Paul Ltd. 1953. Pp. vii + 208. 18*s.* net.

This is the book form of a Ph.D. thesis on a study of subcultures in Great Britain. The author, from New Zealand, attempts to give concrete form to the hypotheses put forward by Kardiner, Linton and others of a basic personality type and the existence of subcultures. It is assumed that if there are subcultures, very different typical personality structures will be apparent in each subculture as well as a basic personality type common to the culture group as a whole. Two very dissimilar subcultures are chosen and the method is to compare the development of children in the two groups, investigating whether different basic experiences make for different personality structures.

Dr Spinley's material for the study of the deprived group, a London slum area, makes fascinating reading for those engaged in any branch of social service work. She observes the personal psychological factors influencing the children—often being unwanted, lack of training in cleanliness, being quickly displaced in the mothers' care by another child, etc. It raises far-reaching practical questions of how best to deal with the delinquent and anti-social products of such an environment, by individual methods or by large-scale social medicine stressing different infant and child training and so aiming at altering the subculture personality pattern.

After the thorough treatment of the slum group, the study of the privileged or Public School group is disappointing and unconvincing. Dr Spinley is aware that much of her information on this group is second-hand and liable to distortion and merely hopes that in some areas of personality enough would be revealed to enable a comparison to be made. This section of the book shows clearly the difficulties and 'pitfalls' of culture studies. It is, for instance, assumed that the boys and girls attending Public Schools come from the same cultural group, but it is very doubtful whether many of the sisters of boys attending the best known Public Schools attend the rather different counterpart for girls. On the other hand, such important factors as being wanted, very early cleanliness training, the strong position of the father in the family, etc., do emerge.

The core of the investigation is a double comparison. From the data on early training, etc., she infers a personality pattern for each subculture group and then measures this against actual Rorschach records of adolescents representative of the groups. Secondly, the Rorschach records of the two groups are compared directly together. Briefly, the records bear out the subculture personalities inferred from the case material and the author tentatively points to certain trends which may belong to the British personality structure. Her conclusions about the latter are weakened by reason of the fact that a large percentage of the slum group is of Irish immigrant extraction and the relevance of an Irish culture seems to have been ignored.

Other criticisms of method might be made beyond the admitted shortcomings of the Public School group study but this is a pioneer work and both the omissions and its merits should be valuable in guiding and stimulating further experimental studies in this socio-psychological field.

M. E. JONES

Adolescence and Youth. By PAUL H. LANDIS. Second edition. New York, Toronto and London: McGraw-Hill. 1952. Pp. xii + 461. 42*s.* 6*d.*

The subject of this book is development between the ages of 12 and 24. (The period 12–16 is called adolescence, that from 16 to 24 youth.) The author is a sociologist and his thesis is that the adjustment problems of this period are posed mainly by the structure and climate of the society in which young people grow up, and that the changes associated with puberty are not, by comparison, of much importance. He does not go so far as to say that adolescence is purely a social phenomenon but he is much more interested in the effect of environmental pressures which he discusses very thoroughly. The author finds that the situations are very different for those living in large cities, for those in small towns and again for those in the remoter country districts in which he is particularly interested. He gives a clear account of the most common problems which young Americans have to solve, grouped under the three main headings of attaining adult status in moral development, achieving 'marital adulthood', and becoming economically independent, showing how they differ in the different environments. The treatment of pubertal development and of the psychological reactions to the various social pressures is (notwithstanding the book's subtitle *The process of maturing*) brief and rather superficial. The book is well organized and well produced with numerous pie-diagrams, bar diagrams and other pictographic devices for presenting the social statistics, and the results of opinion surveys on which the author largely relies.

The chief interest of this work for British psychologists will lie in the obvious similarities and differences between growing up in Great Britain and in these three different American environments and the extent to which they can be correlated with differences in social structures and influences. A.H.D.T.

Parent and Child. By JAMES H. S. BOSSARD. Philadelphia: University of Pennsylvania Press. 1953. Pp. x+310. 40*s*. net.

This book is not a comprehensive treatise on family life and relations between parents and children. Nor is it a contribution to knowledge as its subtitle 'Studies in Family Behavior' might imply: the studies are on the whole too slight for that. But if they are considered as illustrative of the text and as worked examples of suitable field exercises for the beginner, then the book will serve as a useful introduction to the sociology of the family. Its text is readily intelligible (the paragraph on pp. 105–6 is exceptional), it is interesting, especially its studies of the effect of family size, of remarriage, and of differences in class and age between parents upon family life and child development; and it prompts the reader to find out more for himself. H.E.O.J.

The Psychology of Personality. By BERNARD NOTCUTT. London: Methuen and Co. Ltd. 1953. Pp. 259. 21*s*.

The author's preface to this book describes its purpose as that of 'direction finding in the field of Personality'. It is not a book about case histories, nor a technical manual of personality assessment, nor is it a résumé of research findings. The purpose is more to evaluate than to describe the activities of psychologists.

To achieve this purpose the late Prof. Notcutt begins by sorting theories about personality into three broad groups: those using some concept of basic traits (including factors or types), those describing a person in terms of environmental pressures, and those describing the interaction of person and environment. Under this last heading he devotes a chapter to theories of basic motives, which leads into one on patterns of maturation and another on the potentialities of psycho-dynamics. There follow three chapters entitled 'Projective methods', 'Men in groups', and 'The logic of validation' which are in effect brief considerations of methods of personality assessment. In the last seventeen pages Prof. Notcutt tried to outline historically the development of the concept of personality.

Prof. Notcutt has packed a great deal into 255 pages of short sentences remarkably lacking in jargon. Some of his antiseptic cutting and thrusting into the jungle of big trees, little trees, undergrowth and scrub which the psychology of personality resembles today is efficient. Students dazed by impending examinations may thankfully extract some of his tabular arrangements of apparently rival views.

But at this stage praise must give way to criticism. It seems likely that different psychologists would have wanted to take issue with Prof. Notcutt on different points. The reviewer would have wanted a more intelligible and more critical estimate of Lewin and to have asked Prof. Notcutt to go more fully into what he meant by spatial thinking. What he says suggests it is less of a virtue than he supposes. It is also hard to see for what readers the chapters covering personality assessment and history would be adequate. There is a second and more serious kind of criticism to be made. The book sets out to be evaluative, but in spite of its general lucidity it strikes the reviewer as so uneven in its insight, and to reflect such mixed attitudes to the material with which it deals, that one would hesitate to put it in the hands of a new reader without warning. Possibly only a student who is mature and critical enough to take the book lightly would be able to see how useful it might have been but for Prof. Notcutt's odd mixture of scathingly valid and of superficial comment.

JOAN WYNN REEVES

The Forgotten Language. By ERICH FROMM. London: Gollancz. 1952. Pp. 224. 16*s*. net.

An Introduction to Jung's Psychology. By FRIEDA FORDHAM. London: Penguin Books. 1953. Pp. 128. 2*s*. net.

The Psychology and Psychotherapy of Otto Rank. By FAY B. KARPF. New York: Philosophical Library. 1953. Pp. xii + 129. $3.00.

The Forgotten Language is described as a first primer (for everyone) in dream interpretation, and as an introduction to the understanding of dreams, fairy tales and myths. It is based on the view that symbolic language is a language in its own right, in fact, the only universal language the human race has ever developed. As such, it is to be understood, says the author, rather than interpreted as if one were dealing with an artificially manufactured secret code. Dreams and myths are important communications from ourselves to ourselves, and we must understand the language in which they are written.

An Introduction to Jung's Psychology is an exposition for the general reader of the theories and technique of 'the foremost living medical psychologist'. It has an able explanatory introduction by Prof. C. A. Mace, a foreword by Dr C. G. Jung, and a biographical sketch of Jung. It is a clear and sympathetic exposition of Jung's psychology, and will be a useful addition to the books which are aimed at making his ideas intelligible.

The Psychology and Psychotherapy of Otto Rank gives an historical and comparative introduction to the theory and therapy of Rank. His relations to Freud, Jung and Adler and to significant developments in the fields of analysis, psychotherapy, counselling, education and social work, are explained. It contains the well-known photograph of the Freudian inner-circle, taken about 1922, as a frontispiece. Rank appears in this. There is also an interesting photograph of Rank taken in 1924 about the time of his separation from the circle after the publication of his book *The Trauma of Birth*, and another taken before his death in 1939. The present book was begun with Rank's co-operation, and has now been completed and published in response to many requests from interested persons. It will be valuable to have this account of Rank's personality, position in the psychoanalytic movement, and influence today.

R. W. PICKFORD

The Psycho-Analysis of the Nursery. By ALICE BALINT. London: Routledge and Kegan Paul Ltd. 1953. Pp. viii + 149. 12*s*. 6*d*. net.

This book was published in Hungary in 1931 and subsequently translated into German, French and Spanish. But it is only now that it has been possible to produce an English translation.

It contains a clear and interesting statement of some of the most important contributions that psychoanalytic study has made to our fuller understanding of the emotional life of the young child. It describes in considerable detail the strains and problems which the child has to face in his efforts to come to terms with his environment, and the text is enriched and elucidated by examples of children's behaviour and conversation.

One could wish that this book had appeared in England some twenty years ago when the findings would have appeared less familiar and thus more thought provoking. To workers in the professional field it does not perhaps tell anything very new but it serves to reinforce and support our current views on child development. To educators and intelligent parents who already have a background of knowledge of child psychology it will prove stimulating reading, but it is not a book that will be read easily or without considerable hostility by the ordinary parent unaccustomed to such theoretical and often unpalatable concepts as the Oedipus and the Castration Complexes.

This book makes a strong plea for greater honesty of adults towards themselves and towards their children, for an effort to recollect what they knew as children but have in large measure repressed, and to allow the child to share some of the grown-up's feelings. It would be valuable if in addition a simple non-technical statement could be made in a subsequent publication showing more clearly how the knowledge gained from psychoanalysis can be applied in a positive way to the up-bringing of children. On reading this book one is left with the feeling that childhood is a very painful, frustrating and unhappy period when satisfactions are constantly postponed or renounced altogether. Insufficient emphasis is given to the relief from emotional conflict which the child obtains through play, through creative outlets, through acquirement of skills and gradually from new personal relations. It is clear that the translator has not been in close touch with modern developments in English education which allow greater expression of instinctual energy and are less suppressive or authoritarian than would appear in middle class families in the thirties in Hungary.

The translation appears very satisfactory; the style is pleasant and the choice of words makes for clarity and conciseness.

A.H.B.

Developments in Psycho-Analysis. By MELANIE KLEIN, PAULA HEIMANN, SUSAN ISAACS and edited by JOAN RIVIÈRE. London: The Hogarth Press and The Institute of Psycho-Analysis. No. 43 in the International Psycho-Analytical Library, edited by Ernest Jones. 1952. Pp. viii + 368. 30*s*. net.

First Contributions to Psycho-Analysis. By SÁNDOR FERENCZI. London: The Hogarth Press and The Institute of Psycho-Analysis. No. 45 in the International Psycho-Analytical Library, edited by Ernest Jones (who was also the authorized translator of this book). 1953. Pp. 338. 30*s*. net.

The first of these books consists of a general introduction and a chapter on the genesis of psychical conflict in earliest infancy by Joan Rivière; followed by chapters on the nature and function of phantasy by Susan Isaacs; on certain functions of introjection and projection in early infancy by Paula Heimann; on regression by Paula Heimann and Susan Isaacs together; on theoretical conclusions about the emotional life of the infant, on observing the behaviour of young infants, on anxiety and guilt, and on some schizoid mechanisms, by Melanie Klein; and finally a chapter on the life and death instincts by Paula Heimann. There is a useful bibliography. An excellent index has been compiled by Judith Fay.

The book by Ferenczi was originally published in America under the title *Contributions to Psycho-Analysis* (later called *Sex in Psycho-Analysis*), and the present title has been chosen to avoid confusion with other collections of Ferenczi's essays. It is most satisfactory that this valuable book should be available now in the International Psycho-Analytical Library.

No space is available here for an adequate consideration of either of these books. Ferenczi's contributions to psychoanalysis are well known and have been duly estimated and appraised many years ago. The contributions of the school of Isaacs, Klein and others of the same line of thought, are still, unfortunately, in the reviewer's opinion, seriously underestimated in many psychoanalytic quarters, and the publication of this volume should help to clarify some issues, though it is unlikely to close many of the existing splits. It should be read and re-read by every person professionally interested in psychoanalysis and child psychology.

R. W. PICKFORD

The Threshold of the Abnormal. A Basic Survey of Psychopathology. By WERNER WOLFF. London: Medical Publications Ltd. 1952. Pp. xv + 473. 30*s*. net.

Clinical Psychiatry for Practitioners and Students. By IAN SKOTTOWE. London: Eyre and Spottiswoode, on behalf of 'The Practitioner', and in 'The Practitioner' Series. 1953. Pp. x + 395. 36*s*. net.

These are two useful books on psychopathology and psychiatry. *The Threshold of the Abnormal* is divided into six parts, which deal with the concept of abnormality, physiological, environmental and psychological factors of abnormality, treatment, and the diagnosis and evaluation of personality. The author starts by saying: 'While I am not a psychoanalyst of any creed, it is to the memory of Sigmund Freud, father of modern dynamic psychopathology, that the present survey is humbly dedicated.' There are in all eighteen chapters, each followed by a useful bibliography, and the book is embellished with a series of thirty illustrations, some of which are as fascinating as they are instructive.

Clinical Psychiatry is divided into two parts, the first dealing in five chapters with the principles of clinical psychiatry, and the second giving in fourteen chapters clinical descriptions of psychiatric disorders. There is a list of thirty-five 'key references' which is also intended to act as a guide for further reading. This includes psychiatric, psychoanalytic and psychological works. The principal aims of the book are to help the clinician to recognize and deal with psychiatric illness, to know how to examine the patient as a person, what to say to him and to his relatives, what should be done to help him and how to set about doing it.

Both books are to be warmly recommended as useful additions to the existing library of general works on psychiatry and psychopathology.

R. W. PICKFORD

The Psycho-Analysis of Artistic Vision and Hearing. By A. EHRENZWEIG. London: Routledge and Kegan Paul Ltd. 1953. Pp. xxxi+272. 25*s*. net.

This book is, of course, chiefly about vision and hearing as they operate in an aesthetic context. But its interest is greatly increased by the discussion of a number of aspects of everyday perception which are apt to be neglected because neither 'Stimulus'-theory nor Gestalt-theory throws much light upon their character and meaning. Such are the bizarre nature of films shown in reversed temporal direction, the role—constitutive not merely contingent—of roughnesses, transitive tones and other elements which do not appear in the score of a piece of music, and the striking changes in immediate visual appearance which occur as a fashion is born a disturbing novelty, achieves comfortable acceptance and eventually becomes assimilated into past history. Where aesthetic experience or intellectual appreciation are interwoven, the intractable features of perceptual processes become manifold. That official theory is impotent to deal with matters such as these is, in the view of Mr Ehrenzweig, no accident. For it is, at least in part, the function of such theories to conceal from conscious scrutiny the springs of much that moves us in our everyday perceptual experience. This inevitably leads the author into a difficulty—namely that coherent theory cannot provide the linguistic and conceptual tools for discussing what that theory is framed expressly to exclude from consideration. With this difficulty he struggles gallantly, but drags us with him into deep waters, where lurk strange monsters such as 'Gestalt-free structures'. An avowedly psychoanalytic basis assists him not a little in finding appropriate words and ideas. Where these fail he falls back, perhaps a little too readily, upon inverted commas. The book is not easy reading in consequence, but there can be no question of the wealth of interesting and important questions it raises and discusses. It is not easy to agree with all the author has to say. Not every biologically-minded reader, for instance, will share Mr Ehrenzweig's conviction (upon which he leans heavily) that a grave crisis in human history occurred when the adoption of the erect posture first concealed the female *genitalia* from view. Nevertheless, we must be grateful for his originality and courage in tackling so many problems in perception about which many psychologists adopt a defeatist attitude. R.C.O.

Tragedy and the Paradox of the Fortunate Fall. By HERBERT WEISINGER. London: Routledge and Kegan Paul Ltd. 1953. Pp. 300. 21*s*. net.

The author believes that the pleasure we take in tragedy comes from our responding to the fall of the hero as if it were equivalent to the death of the god or divine king in the myth of rebirth. This myth he believes to be an archetype, and therefore 'in responding to the concept of *felix culpa*, we are at the same time responding to all the layers of association embedded in the myth and ritual pattern in the long course of its development'. Unfortunately he offers no coherent, let alone cogent, argument for his belief. In place of this he traces, in a long, dragging catalogue, the development of ritual and belief centring in the death and rebirth of the divine king in the ancient Near East, in the Mediterranean area, and in Hebrew and Christian thought. There is a brief discussion of some modern theories of tragedy, a long and barely relevant attack on the Motion Picture Production Code and a disquisition on the concept of freedom for the arts and literature, with examples of prosecutions in the United States. Although the author's central theme might conceivably be argued more lucidly and convincingly, and to that extent is worth attention, the book as it stands is unprofitable. It is also disfigured by slovenly writing and spattered with such sentences as '...he sympathized with that attempt as only one who followed it in all its bitter and frustrating detail alone would'; '...it is quite clear that they did not, nor could they have had, any effect on the making of the Western mind'; 'suppose...that this assemblage then lays down the rules of the practice of these crafts, obedience to which the pilots and doctors are to be tested annually in open court'. And there are three mistakes in the few lines quoted from Milton on page 19. This degree of literacy (or care) falls short of what is needed for the writing of a book. D.W.H.

Animals and Men. By DAVID KATZ. London: Penguin Books. Revised edition. 1953. Pp. 192. 2*s*. net.

In the preface to the original edition of *Animals and Men* (1937), the author explains that he intends to present a 'few problems of animal psychology on which I worked with my students... (and) which contribute directly to comparative psychology'. This preface had been omitted in the new revised edition, and the reader is left with the unfortunate impression that this slim volume forms a representative survey of animal psychology.

Except for references to the Lorenz-Tinbergen school, little new material has been added. The reader will not find adequate reports of the intensive studies of subprimates at the Yerkes Laboratory or

Wisconsin, of the investigation of the behaviour of honey-bees by von Frisch, of the transplantation experiments of Weiss and Sperry, the work on orientation of lower animals previously reviewed by Frankel and Gunn (1940), nor of the innumerable experiments on the laboratory rat—special appetites, maze learning, brain mechanisms in instinctive and learned behaviour, etc.

Furthermore, some of the sections which should have undergone extensive revision have remained untouched. For instance, the discussion of homing and migration of birds makes no reference to the contributions of Griffin, Bissonette or Bullough, and the description of Koehler's insight studies, using chimpanzees, does not mention the follow up studies of Birch (1945).

Despite these omissions the book contains many interesting references to Katz's own extensive work in animal psychology, mainly with hens and monkeys, and surveys a considerable number of experiments and field observations by European workers up to the outbreak of the last war. Katz's general orientation is avowedly phenomenological; the method demands that each piece of research is undertaken with an unbiased description of the objects to be studied. Consequently, there are few excursions into theory, and except in the chapter on 'Needs, Drives and Instincts', theoretical problems are by-passed. We are handed certain unresolved puzzles, like homing and migration in birds, the homing of dogs, and counting in birds and monkeys, but are not taken behind the scenes, as it were.

The comparisons between animals and men are not always happy: 'there is a resemblance between the faulty development of wild animals in captivity and the neurotic's inadequate adaptation to his environment', or the refutation of the psychoanalytic view that thumb sucking is an expression of infantile sexuality by reference to Levy's experiments with puppies.

The present translation by H. Steinberg and A. Summerfield is excellent, and makes fluent reading. The book would have gained in its usefulness to psychologists had more of the original references been retained. However, it deserves to be widely read by those wanting a stimulating introduction to animal psychology.

H.M.B.H.

The New Way to Better Hearing. By V. L. BROWD. London: Faber and Faber. 1953. Pp. xii + 252. 12*s.* 6*d.*

This book, by an American otologist, is addressed to two broad groups of readers: (1) those suffering from defective hearing, and (2) those concerned with their rehabilitation. It presents a non-technical account of the author's system of hearing re-education, together with chapters on the ear and hearing, and detailed proposals for a course of auditory training.

Dr Browd points out that in almost every case of deafness there is present some residual hearing capacity which is often not used by the patient. He bases his system upon training the ear to utilize this residual capacity to the full, particularly in the discrimination of speech sounds. Unlike some authorities he discounts specific training in lip-reading as 'preventing the freeing and conversion of hidden hearing power'. The detailed course of training requires the services of an untrained assistant, and takes up ten minutes a day.

The author's emphasis on perceptual processes in auditory training is stimulating and salutary, and much of his practical advice is sound. The book should be particularly valuable to a hard-of-hearing adult who is having difficulty in understanding speech or in using a hearing aid. One must remain sceptical about the value of the system with profoundly deaf patients: one frequently gains the impression from the book that in such cases useful hearing for speech can be restored. No experimental evidence is produced in support of this claim, although it is backed up by a number of anecdotal case histories. Dr Browd recognizes the need for social rehabilitation for the deaf, but his remedies may seem somewhat aggressive to English readers, and to smack of 'bookstall psychology'. In one of his demonstrations of hearing at increased distances he appears to ignore the fact that the inverse square law does not apply within an enclosed space of the kind he describes.

The book has neither index nor references. There are, however, eight excellent colour plates (very much better than the illustrations of the ear found in current psychological text-books) and its price is reasonable.

D. C. KENDALL

A Guide to Mental Testing. By RAYMOND B. CATTELL. London: University of London Press. Third edition. 1953. Pp. xv + 446. 35*s.* net.

This book was first published in 1936 and is too well known to require a detailed introduction. In conformity with the author's intention of revising, 'at intervals in response to the progress of research and the publication of new tests', a second edition appeared in 1948. The present enlarged and revised version contains a number of minor changes and a few major ones. Anachronisms which remained in the last edition have been removed. For example, the reference to 'Cox's recent research' into tests of manual dexterity has been amended to read, 'Cox's basic research'. The chapter on the Measurement of Intelligence now contains new and more conveniently presented norms on the Cattell Intelligence Scales.

The major changes occur in the chapters on Tests of Temperament and Disposition, and Measures of Character Factors and Probes of Emotional Adjustment. In the first of these a Brief Selective-Answer Misperceptive Test of Personality Factors has been included. It has been constructed to measure, 'five of the more important and firmly established primary personality factors'. Two other new tests of personality have also been included. One is a humour test and the other a musical preference test. Both are reported to distinguish normal from abnormal populations.

In the second of the chapters mentioned above Cattell discusses his Sixteen Personality Factor Questionnaire and includes the items from the shortened form of this test. This is apparently intended to replace the Questionnaire for Personality Factors which was published as Appendix II in the last edition.

Among the tests that users of this guide might wish to see mentioned is likely to be the Wechsler Intelligence Scale for Children. Although this test is becoming increasingly popular no reference is made to it. It is perhaps unfortunate that the theoretical limitations that the author imposes upon himself are passed on to the practising psychologist in this way. However, if these limitations are accepted, *A Guide to Mental Testing* provides a convenient reference to a wide range of useful tests. A. RICHARDSON

The Fourth Mental Measurements Yearbook. Edited by O. K. BUROS. Highland Park, New Jersey: Gryphon Press. 1953. Pp. xxvi+1163. $18.00.

This reference work is designed to help test users to 'locate and evaluate tests and books on testing' (p. xix). It covers the period 1948–1950 and thus complements rather than replaces its predecessor *The Third Mental Measurements Yearbook* (covering 1940–1947). Each test is named and numbered, information is given about the source of supply and cost on the one hand, and about standardization, reliability and validity on the other. This is followed by evaluative reviews written specially for the *Yearbook*, together with excerpts from reviews published elsewhere. Comprehensive lists of book and journal references to the test are also included. The main part of the book (830 pages) is devoted to these test entries, the tests being classified on the system used in the earlier *Yearbook* with only minor modifications. The remainder of the book deals with reviews of books relating to mental testing, directories of periodicals and publishers, and three indices: names, titles and tests.

Dr Buros has again provided psychologists, educationalists and other test users with a most valuable reference book. He has, moreover, performed an additional service that becomes apparent immediately one starts to use this book or its predecessor. In the detailed instructions given to the contributing test reviewers Dr Buros states his aim: 'to stimulate progress towards higher professional standards in the construction and use of tests, by commending good work, by censuring poor work, and by suggesting improvements'. This editorial policy, supported by more than three hundred special reviewers, makes the book valuable not only as a work of reference but also as a very thorough piece of 'reality testing' of one of psychology's most important tools. Good, well-standardized tests are given the credit due to them, but no reader can long remain unaware of just how many tests of unknown reliability or uninvestigated validity are in widespread use. The reviewers have not failed in the task assigned to them of making their criticisms specific, of mentioning alternative better-standardized tests when available, and of indicating weaknesses or omissions of fundamental kinds in test construction and standardization. The book is, as a whole, rather depressing in that it indicates just how widespread faults of this kind are; it also serves to make this fact unmistakably plain to the world of professional psychologists. Those who consult it must inevitably come to sympathize with its editor and his contributors in their efforts 'to impel test authors and publishers to present more detailed information on the construction, validation, uses and misuses of their tests'.

These books, which began with *Educational and Personality Tests of* 1933 *and* 1934, have shown progressive development. Future volumes (which it is hoped to publish every three years) might perhaps develop a stage further to include a separate list of approved, properly standardized tests. The residue of the book could comprise tests of dubious value, together with those at the quite legitimate stages of research development; this residue would form the main body of the book. It is unfortunate that both this and the earlier *Yearbook* are so bulky; this limits their usefulness, since it is a laborious task for each reader to evaluate the better tests for himself. A place on such a 'white list'—one hesitates to suggest a 'black list'—might in the course of time come to be eagerly sought. This would greatly hasten the achievement of the wholly admirable aim of the book's editor and contributors. Separate publication of the section on books as opposed to tests would also diminish size and costs. Again the present system of test classification is not wholly satisfactory; larger printed numbers would assist quicker location of test entries; the Universal Decimal System of numbering merits consideration; and tabulation might be used to shorten the test reviews. Yet it would be ungrateful to press minor criticisms of a major contribution towards the more rigorously scientific developments of psychology. Despite its cost this is a book which no responsible institution concerned with psychological testing can afford to do without.

PETER MCKELLAR

The Tools of Social Science. By JOHN MADGE. London: Longmans, Green and Co. 1953. Pp. x+308. 25*s*. net.

Methods grow out of problems and a periodic evaluation of the techniques for the collection and analysis of data is a necessary part of any developing science. This evaluation forms the main section of Madge's book and provides at the same time a well-presented introduction to such topics as the use of documents, observation and interviewing methods, attitude scaling, sampling procedures and the design of experiments in social science.

However, this book is more than its title might suggest, since about a quarter of the total space has been devoted to a discussion of such fundamental theoretical and practical issues as the place of value judgements, the nature of objectivity and the function of language and logic. For example, in discussing value judgements the author writes: 'Instead of aiming to become the objective value free theorist, the social scientist may without shame take on the job of social engineer, in which he may use his experience and special skills in the service of his chosen community.' This is a recurrent theme throughout the book and one which is integrated with the evaluation of the techniques at present available.

'Exactness is beyond our reach because of the penumbra that surrounds all our observations, and universal truth is not possible because every truth that we adhere to contains within itself some part of our aims, and aims are never unanimous....The truth that is powerful in the world is the truth that is sufficiently exact to get things done.' These views may not commend themselves to the purists but, in a period when social science has tended to retreat into the abstractions of theory and technique, they supply a timely counteractive.

A useful source of relevant references is provided by the footnotes which are neither too intimidating nor too distracting. As a whole the book is well presented, the elimination of much technical jargon making it as suitable for the interested layman as for the social science student.

A. RICHARDSON

Appraising Personality. By MOLLY HARROWER. London: Routledge and Kegan Paul Ltd. 1953. Pp. xviii+197. 18*s*. net.

This is an attempt to make known to doctors, psychologists and a wider public something of what takes place when the clinical psychologist assesses personality by projective methods. It is in the form of dialogues between a psychologist and a physician, the latter a patient pupil and a little of a Watson—a way of presentation that proves somewhat trying. However, the exposition is clear, simple and free from jargon. The projective techniques include the Rorschach, drawing of the human figure and of the 'most unpleasant concept', the Szondi, and a promising sentence-completion test. Personality clues elicited by the Wechsler Bellevue test are also used. No attempt is made to cover the techniques of administration, the aim being rather to illustrate the information that can be obtained and the use that can be made of it.

There is a suggestion that the projective responses are regularly striking and impressive, the psychologist always right and the outcome invariably successful: this is perhaps but the result of the attempt to write a popular book on a notably tricky subject. Something of the difficulties and limitations as well as the conspicuous successes of the techniques would have given a more realistic picture and, probably, added to the reader's confidence in the methods discussed. However, the author's handling of the projective material is shrewd and clearly rests on sound experience.

The Yearbook of Psychoanalysis. Vol. VIII. Edited by SANDOR LORAND. New York: International Universities Press. 1953. Pp. 383. $7.50.

This volume of *The Yearbook of Psychoanalysis* collects re-printed contributions by twenty-five authors, including Freud himself and many well-known writers on psychoanalysis and related subjects. It is full of interest for the technical reader and many of its contents will command the attention of the general psychological reader as well. It maintains the high standard and general interest of previous volumes and will be a stimulating book for addition to any library containing works of psychoanalysis.

R. W. PICKFORD

Proceedings of the Fourth International Congress on Mental Health, Mexico City, December, 1951. Mexico: La Prensa Medica Mexicana. London: H. K. Lewis and Co. Ltd. New York: Columbia University Press. 1952. Pp. xvi+386. Eight photographs of the Congress. Four photographs of murals by Guillermo Meza. 30*s.* net.

To all psychiatrists, psychotherapists, psychologists, social workers and others professionally interested in mental health, this volume of the proceedings of the Fourth International Mental Health Congress will be of the keenest interest. It is edited by Alfonso Millan, and contains, in English, a complete record of the work of the Congress, which covers every aspect of the problems of mental health, psychiatric, psychological, psychoanalytic, sociological, educational and racial. R. W. PICKFORD

Catalogue of Films on Psychology 1953. London: The Scientific Film Association. 1953. Pp. vii+118. 7*s.* 6*d.*

This is an extremely valuable publication. As a catalogue it is not quite complete. There are, for instance, a number of films on animal learning and child development held by the University of St Andrews which are not included; but this is a defect which will easily be remedied now that a first step has been made, since all those departments who have films will no doubt make a point of having them listed.

What gives this catalogue its value is the short appraisal given for each film. The reviewer has checked the appraisals of those films he has seen and finds himself in complete agreement with them.

It is perhaps unfortunate that so many of the films that have been made in the field of psychology are of a rather superficial kind and directed towards a general audience. It may be that the expenses involved make the production of more technical films difficult, or it may be that the film itself is not a suitable vehicle for the conveying of complex information. This does not, however, detract from the value of the job done by those who have compiled this catalogue which should be in the hands of everyone running courses in psychology, education or social sciences. JAMES DREVER

CORRIGENDUM

It is regretted that the author of *The Ames Demonstration in Perception* reviewed in Vol. XLIV, p. 268, was given incorrectly. It should be William H. Ittelson.

VOLUME XLV AUGUST 1954 PART 3

EXPECTANCY IN APPARENT VISUAL MOVEMENT

BY E. E. JONES AND J. S. BRUNER
Harvard University

I. *Introduction* (p. 157). II. *Method* (p. 158). III. *Experiment I: procedure and results* (pp. 157–158). IV. *Experiment II: procedure and results* (pp. 159–161). V. *Experiment III: procedure and results* (pp. 161–163). VI. *Experiment IV: procedure and results* (pp. 163–165). VII. *Discussion* (p. 165). *References* (p. 165).

I. INTRODUCTION

The person responding to stroboscopic movement in the experimental room is familiar with a plethora of moving objects in a great variety of meaningful contexts before he ever comes to the experiment. What difference does this make? Out of familiarity comes a matrix of expectancies ordering and interrelating the cues to apparent movement. The perception of movement and its various attributes may be conditioned by these previously established perceptual expectancies. Put simply, the organism learns to expect movement in some contexts more than in others, and learns to expect certain *kinds* of movement from some objects and not from others.

The present series of experiments seeks to demonstrate some of the ways in which movement-related expectancies condition the perception of movement in a typical stroboscopic experiment. Expectancies were aroused by manipulating the nature of the objects in motion and the stimulus context in which movement occurred. The first two experiments deal with comparisons of the properties of apparent movement in 'meaningful' and 'meaningless' contexts. The second two are more concerned with the promotion of different patterns of movement when spatio-temporal factors are held constant, but the stimulus context and consequently the psychological meaning of movement is altered.

The importance of conditions related to past experience has been recognized by many students of movement phenomena, but most research on apparent movement has reflected an interest in the defining parameters of movement such as the interdependent spatial, intensity, and temporal factors. Wertheimer's early remarks suggesting the temporal intervals appropriate for simultaneity, optimal movement and succession, and the four laws of parametric interrelationship propounded by Korte, have been modified and extended by later research (see Graham in Stevens, 1951). Willey (1936) has demonstrated consistent individual differences in the resolution of ambiguous movement patterns. Duncker's (see Ellis, 1938*a*) exciting monograph on 'induced movement' opened a broad area concerning the relation between a moving object and its framework. But, by and large, research and theory in this area have been directed toward analysis of the *necessary and sufficient* conditions for producing stroboscopic movement. While such research must come first, it is also appropriate to extend the analysis of movement by attending to more cognitive determinants of the movement experience. It does not suffice merely to hold these constant—'meaning' held constant is indicative of the fact that we choose to overlook a large source of variance.

II. METHOD

In the experiments to be described below, eighteen Harvard students were used as subjects in Exps. I and II, while seventeen of these served in Exps. III and IV. They went through the experiments individually and in one sitting. The stimuli for apparent movement were projected upon a white cardboard screen held in a frame 22 in. high and 29 in. wide. The subjects sat at a slight angle to the screen, approximately 8 ft. from it. Slightly to the left of the subjects, directly facing the screen and $8\frac{1}{2}$ ft. from it, was a table on which were placed three Besler OA episcopes in such a way that the light projected by each covered an identical area coinciding with the limits of the white screen.

The three episcopes were wired to a timer regulating the exposure duration for each projector as well as the interval between exposures. The timer consisted of a Bodine Synchronous Motor driving a rotating shaft on which were mounted three pairs of circular steel disks with variable notches in their circumference activating micro-switches. The width of the notch in a given disk (wired to a given episcope) determined the exposure time for the episcope. The position of the variable notch on the disk determined the time interval elapsing between exposure of any two episcopes. The experimenter operated a simple control panel of three noiseless switches that opened or closed the circuit between the timer disks and each episcope. Movement could be produced by either two or three successive exposures depending on how many circuits were open.

In each episcope was placed a stimulus card on which pictorial material was drawn. The pictures used in the present experiment were all drawn in indian ink on white cards superimposed on larger black cardboard. The projected size of the white surface containing the stimulus pictures was, as we have said, coincident with the size of the white screen so that subjects were not distracted by nearby contours of the room, which was completely dark except for the episcopes in exposure.

Under these conditions, movement was quite easily created by the successive exposure of each episcope. The range of optimal movement was wide for any given subject. In the majority of the situations, movement could be maintained from as close to simultaneity as 0·1 sec. interval separation to as close to succession as 0·4 sec.

III. EXPERIMENT I: PROCEDURE AND RESULTS

In the first experiment subjects were shown the sequence of cards presented in Fig. 1. In the successive exposure of the stimulus cards the 'stick-man' moved from the left side in the first phase to the right side in the third phase, changing places with the nonsense object which moved an equal distance in the opposite direction. Each subject was shown the picture sequence for one trial, asked to report his perception, given four more trials, and asked for a second report. Reports were spontaneous descriptions; no directive instructions were given. On each trial, the stimulus cards were separated by an interval of 0·15 sec., the exposure time being 0·20 sec. for each episcope. If a subject saw both objects in motion, he was asked to compare the objects with regard to speed of movement and relative distance travelled.

The results need no statistical elaboration. After the first trial, fifteen of the eighteen subjects reported seeing the man as moving faster than the nonsense object, and sixteen of the eighteen reported that the man covered a greater distance in his movement. Both results are at the 0·01 level, computed by the Dixon & Mood (1946) sign test. One of the most striking findings was that on the first trial, six subjects (or one-third of the cases) saw only the man as moving and in one case the nonsense object was not even seen.

On the fifth trial this relationship no longer obtains. Eight of the subjects saw the man and the object moving at an equal speed, and eight of the remaining ten saw the nonsense object as moving faster. This same shift occurs with respect to the perception of distance travelled. We may conclude that the results are not simply a function of constant position factors—that left-to-right movement is somehow more compelling than right-to-left movement. A more reasonable conclusion is that a man-in-motion has a higher subjective probability value than a thing-in-motion. This expectancy operates to accentuate the perceived velocity and distance of the man's movement relative to that of the nonsense object on the first trial when the actual spatial-temporal features are still equivocal, and when the objects are not yet clearly localized within a stable framework. When the framework is established, there appears to be a tendency toward reversal. We do not know how to account for this effect save to suggest that Köhler-Wallach (1944) satiation may be operative in apparent movement.

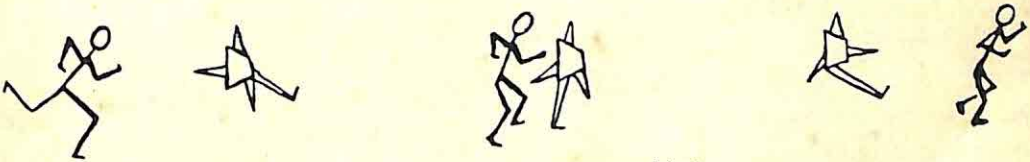


Fig. 1. Stick-man and nonsense objects.

IV. EXPERIMENT II: PROCEDURE AND RESULTS

In the second experiment we turn to another way in which a meaningful relationship between subjective expectancy and context may accentuate apparent movement. Optimal stroboscopic movement can only be maintained within certain limits of spatial and temporal separation between stimuli. One might expect these limits to be broader when the context in which the movement occurs is one in which motion has been frequently experienced. In other words, movement in an often experienced context would be expected to maintain itself more effectively under suboptimal conditions than movement without such experiential support.

Fig. 2 shows the sequence of cards used in the present experiment.

All of the eighteen subjects were initially exposed to thirty trials of both the Control and Experimental Series, ten trials each at separation intervals of 0·15, 0·20 and 0·25 sec. Then each subject was further exposed to fifteen successive presentations of the control series (oblong-over-arc stimulus) and fifteen successive presentations of the experimental series (car-over-bridge stimulus). The exposure for each card in the two series was 0·20 sec., with an interval between cards of 0·15 sec. Half of the subjects received the experimental series first and half of the subjects received the control series first. On four randomly interspersed trials in each series, the middle stage of movement was omitted by cutting the current to the centre episcope. Thus the objective stimulus consisted merely of a car at one end of the bridge and then a car at the other end, or the nonsense version of the same two exposures.

In the control series of fifteen trials, two-phase exposures occurred on trials 4, 7, 10 and 11. In the experimental series, two-phase exposures occurred on trials 5, 8, 11 and 12.

The subjects were told: 'I'm going to show you the same pictures in a series of fifteen consecutive trials. I want you to look at the total series to judge whether there seems to be any progressive trend in goodness of movement from trial 1 to trial 15. In other words, does it seem to be getting consistently faster, or slower, or going farther or less far, or seem smoother or jerkier? I want you to report specific fluctuations from trial to trial—as you see them—on any or all of these characteristics.'

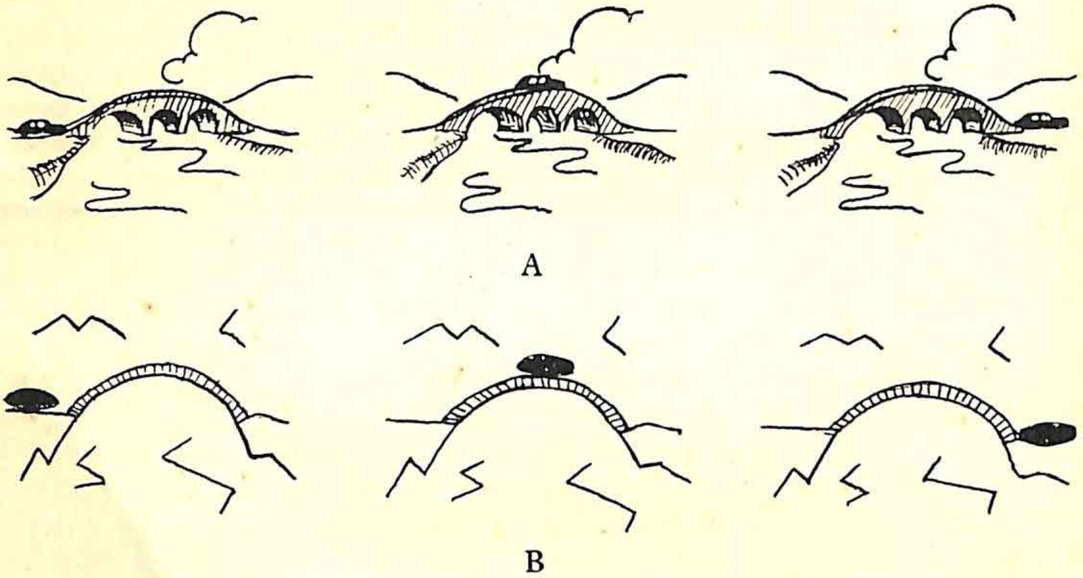


Fig. 2. A, car-over-bridge context; B, oblong-over-arc context

Table 1. *Responses to two-exposure trials for experimental* (*auto*) *and control* (*oblong*) *series*

	Control		Experimental	
No change in character of movement		36		57
Change in character of movement				
(*a*) Unevenness noted	15	36	5	15
(*b*) Middle image missing	21		10	
Total number of responses*		72		72

* This total represents four responses for each of the eighteen subjects.

On the basis of differences in strength of movement-expectancy, we would predict more disruption of movement on critical two-phase trials under nonsense conditions. That is, though the expectation of three-phase movement is reinforced in the experimental situation for both card sequences, movement in the more meaningful context will be *relatively* less affected when the mid-phase is absent, since the expectancy in the latter case is well buttressed by previous confirmations outside of the experimental situation. Table 1 classifies the spontaneous response to the trials in which the middle exposure was omitted.

Without the benefit of statistical test[1] these differences seem rather impressive by inspection: 79% of the observations under the experimental condition were reports of no change, whereas only 50% of the control reports were 'no change' responses; only 14% of the responses under experimental conditions indicated perception of the missing middle image, whereas 29% of the control responses did so.

---

[1] Inadvertently the original data for this experiment were destroyed. The combined data in Table 1 do not permit the employment of such usual tests of contingency as χ^2.

To the extent that these differences are reliable, we conclude that the meaningful context more effectively mediates the expectancy than does the nonsense situation. The more the stimulus context of movement is symbolically appropriate, the less susceptible will the experience of movement be to disruption by a change in physical conditions away from the optimal state—such as a reduction in the number of exposures in a given movement space.

V. EXPERIMENT III: PROCEDURE AND RESULTS

In the present experiment our object was to study the patterning of movement as a function of the intrinsic characteristics of the field. Fig. 3 contains the stimuli used. In one version of the three-phase movement, the ball on the inclined plane is seen to *skid*—a movement mediated by the constancy of position of the inset arc. In the other version, the ball is seen to *roll* down the inclined plane—the arc changing position appropriately in the three exposures. Movement characteristics should be appropriate to each context—a skidding context and a rolling context. Thus, apparent speed and distance traversed should be greater in the second instance.

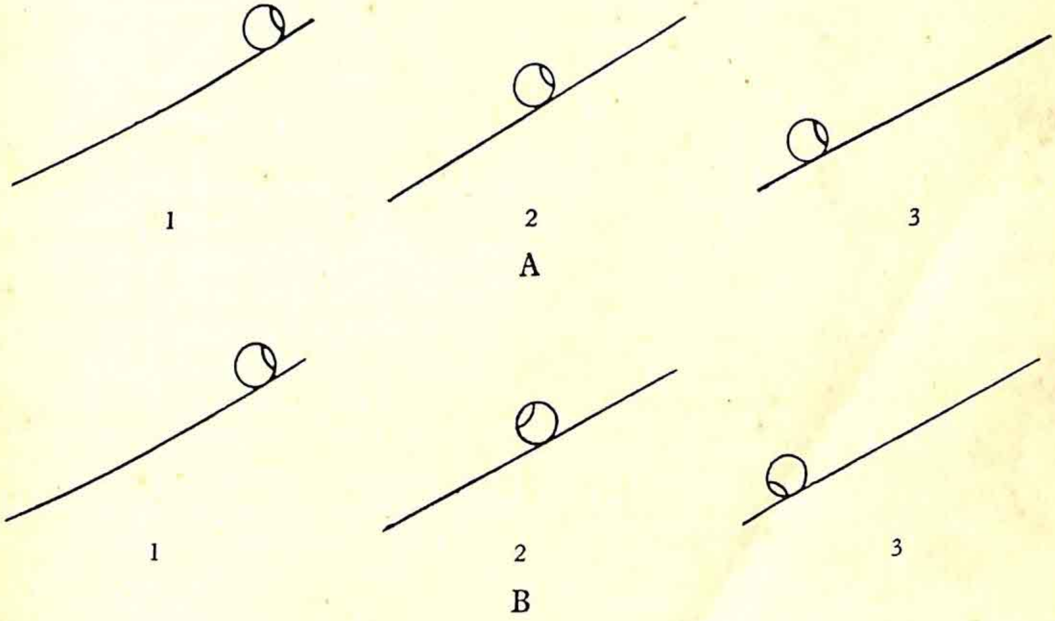


Fig. 3. A, 'skidding' ball context; B, 'rolling' ball context.

Each of seventeen subjects was shown the 'skidding' ball (Fig. 3A) for ten trials at an exposure time of 0·20 sec. and a separation-interval of 0·15 sec. They were then shown the 'rolling' ball (Fig. 3B) for ten trials and asked to compare the two sets of exposures. Direct questions were not asked for fear that they be leading questions, but the subjects were asked to describe what they saw, and an attempt was made to classify their spontaneous responses.

Eleven of the seventeen subjects saw the ball in Fig. 3A as skidding, none of them saw it consistently as rolling down the plane, and the remaining six subjects saw it in a reverse roll or in some other fashion. Where the ball in Fig. 3B was concerned, twelve subjects saw it as rolling, none as skidding, and the five others gave responses that were not classifiable. The test of our hypothesis comes in the comparison of the two forms of movement. Fifteen of the seventeen subjects saw the 'rolling' ball sequence as faster, thirteen of the seventeen as covering a greater distance (χ^2 yielding values in both cases at less than the 0·05 level). Quite evidently, the 'rolling' ball was seen to have more velocity and more momentum.

Qualitative distinctions between the two movement experiences were clear. The difference between the two series was described specifically as a difference between skidding and rolling circles by eleven of the subjects, and in all cases the two movement sequences resulted in qualitatively distinct response patterns. Some of the spontaneous comments of the subjects are presented below to indicate the variety of attempts to resolve the tension produced by the circle with the constant inset arc. Many of these attempts involve an active reorganization of objectively constant background features in an attempt to maximize the meaning of the movement and thereby resolve the conflict.

Subject AH watched the skidding sequence and, after one trial, said that the ball was rolling down the hill. After the second trial he said: 'It's still a circle rolling down a hill, but it starts uphill before it goes downhill.' This impression remained until the sixth trial when he said that the movement was very jerky and almost broke up. On the seventh trial he reported: 'This whole picture bothers me. It doesn't move correctly because I have the feeling that it's going backwards—the hill should be facing the other way.' After the first trial of the rolling sequence, AH reported spontaneously that 'the diagonal has shifted and the hill is much steeper now'. After the second trial, he revised his judgement by saying it was not as steep as he had thought, but that it was still steeper than the first sequence. After the fourth trial, the same subject reported that the circle seemed to hesitate at first and start slowly, but when it got started, it gathered a great deal of momentum which made it 'go a long way off the picture at the bottom of the screen'. When asked to summarize his comparison, he stated that the second series of movements was much faster and farther. At no time during the experimental trials was this subject able to locate the essential difference between the two series of the position of the spot. He also never referred to one series as skidding and the other as rolling, and the experimenter made no suggestion in this direction.[1]

Spontaneous observations given by other subjects to the skidding-ball stimulus were: 'The ball is pushed down, it skids to a stop.' 'I see a ball, but I can't tell whether it rolls up the hill or down.' 'The movement is very jerky and the ball stops a long way from the edge of the screen.' 'I'm surprised to see it [apparently the arc] reappear in the same position. You look for it in a different place and you don't see it.' 'It gives me a retarded sense and seems to skid to a stop.' 'The ball isn't rolling, but it's balanced on a string and somebody's pulling the string.' 'A ball is rolling down a wire but something is funny. I can't spot what that extra line is, but the whole thing seems wrong.'

It might be pointed out, as a final interpretative note, that this experiment sheds further light on the concept of *identity* in stroboscopic movement. Linke (see Neff, 1936) was the first to affirm that phenomenal identity of initial and terminal stimuli was the one essential determinant of apparent movement perception. But it was Ternus (see Ellis,

---

[1] The reports of AH argue against explaining the observed modification of perceived movement in too intellectualistic terms. At least in his one case, this response is not a logically considered comparison but an immediate and unreflective response to the total field.

1938*b*) who underlined the ambiguity of the identity concept by demonstrating the complexity of spatial identity in even a simple stimulus situation. Working with four equidistant points lying along a straight line, Ternus presented the first three points as the initial stimulus, and the last three points as the terminal stimulus (so that the two middle points were exposed both times and in the same position). Instead of seeing a single point move from the first to the fourth position, his observers perceived the movement of a line composed of three points. In the movement of such a figure, only homologous parts (parts playing the same role) exhibit phenomenal identity, since 'phenomenal identity is a characteristic not of piece-wise relationships, but of whole-phenomena' (Ellis, 1938*c*). Identity must be appropriate to the pattern of movement.

Orlanski (1940) addressed himself to a different kind of identity—the identity of stimulus form in successive exposures rather than spatial location. Progressively changing the 'identity' of his stimulus objects by varying formal dimension, he showed that the stability of movement (using a range index of stability) varied directly with the objective identity of the stimulus. It is unfortunate that he did not examine more carefully the distinction between identity of stimulus form and identity of stimulus function. It is tautological to say that a moving object must be seen as a moving unity in sequential stages. However, this is not equivalent to saying that, when objects are presented in succession, the more closely alike they are in form, the more compelling will be the movement experience. The present experiment shows that there are stimulus conditions in which movement under conditions of *formal* identity is *less* compelling than movement of a 'changing' object when the change conforms to well-established expectancy and is functional to the requirements of total field conditions.

VI. EXPERIMENT IV: PROCEDURES AND RESULTS

This experiment demonstrates the effects of meaningful context on the option of the path of apparent movement. The general hypothesis is that where there is an option in the path of apparent movement, that path of movement will occur, other things being equal, that most conforms to an established expectancy. The stimulus sequences used in the experiment are presented in Fig. 4. In series A we see a man on the lower left and a man on the lower right, both throwing baseballs diagonally across a space to the top corners. The two balls are adjacent in the middle of the card in the second exposure. In series B there are two billiard players, both of whom are hitting billiard balls from the upper two corners of the pool table, diagonally across to the two lower corners. In the mid-phase of movement the balls are again seen to be adjacent. Series A was shown at a separation-interval of 0·20 sec. and an exposure time of 0·20 sec. per card. After several intervening experiments, series B was presented at the same interval and exposure. Both series were presented for ten trials.

After each trial of each series, the subject was asked to report all the impressions of movement he had received. If he insisted in centering his responses around the 'actors' in either series, the experimenters directed his attention by more specific questions to the course or movement taken by the balls. In all but a few cases, the report situation was unstructured. Responses to each of the ten exposures were recorded for each series. If the situation was not understood at all (as was frequently the case with the pool players), the initial stimulus card was taken from the episcope and shown to the subject.

These particular sequences were chosen in an attempt to overlay a basic pattern of movement ambiguity, described by Ternus (see Ellis 1938*d*), with symbolic contexts which would predispose the subject to choose a specific alternative. Ternus had found that the apparent movement of dots placed as ours were could be seen either as 'crossing over' or as 'cannoning' or 'bouncing' off each other. In designing the stimuli, it seemed to us more probable that two thrown baseballs would not hit each other in mid-air and cannon off at an angle. On the other hand, it seemed very probable that two pool balls might collide and be deflected into opposite corners.

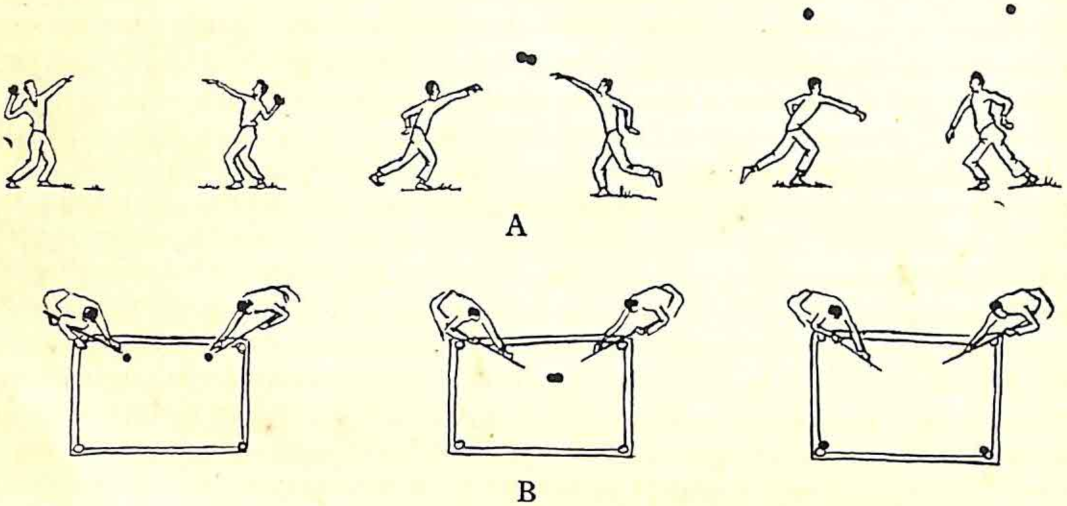


Fig. 4. A, baseball context; B, billiards context.

Table 2.† *Responses to ambiguous movement patterns in two contexts*

	Cross over	Bounce	Total	
Baseball contest	10	5	15	$\chi^2 = 1{\cdot}6667$* $P = 0{\cdot}1998$
Billiard context	4	11	15	$\chi^2 = 3{\cdot}2667$* $P = 0{\cdot}0753$

† Table 2 contains a comparison of *each* context against 50:50 null hypothesis. Direct comparison of the two contexts with each other is not possible because the same subjects were used in each context, thus violating the requirement of independence necessary for direct comparison by the χ^2 technique. Other, less direct techniques, could not be applied because of the inadvertent loss of the raw data previously mentioned.

The results tend to support the prediction but not conclusively. In the majority of reports, the pre-existing expectancy of the subject promoted the predicted option for each sequence. Disregarding the 'ambiguous' and 'no movement' reports of two subjects, we can construct a contingency table (Table 2). The association between context and response option (pooled probability) is nearly significant ($P = 0{\cdot}076$). Thus the same basic 'carriers' of the movement experience may be seen in two qualitatively distinct patterns when different expectancies are involved.

VII. Discussion

This paper has attempted to demonstrate the operation of a variable which conditions the perception of visual movement. The variable of our interest may be called variously stimulus *context* (the experimentally manipulated stimulus framework in which movement occurs), *expectancy* (the pre-established network of hypotheses about movement in the subject), or *meaning* (the relationship between stimulus context and subjective expectancy).

We have cited and briefly described four experiments in which the relationship of stimulus context to postulated general expectancies was shown to be an important source of variation. The evidence from these experiments supported the following hypotheses:

Experiment I. When two objects cover an identical distance objectively, the object which has more frequently been seen in motion will be seen initially as moving faster and farther.

Experiment II. The more probable and practised an apparent movement, the more adequately will the movement experience maintain itself under sub-optimal conditions.

Experiment III. Speed and extent of apparent movement are related to the qualitative form of movement promoted by the stimulus context.

Experiment IV. When the direction of movement is basically optional or ambiguous, there is a tendency for contextual cues to promote one option or another according to the expectancy aroused.

The experiments reported here are of the order of 'demonstration experiments'. It would be a further amplification of the problem to establish experimentally certain expectancies of movement in our subjects and to work with these as independent variables. Our object in reporting the present experiments has been to underline the importance of such cognitive variables as expectancy in the determination of speed, extent, path, and consistency of apparent movement. Research on the parameters of apparent movement perception has generally neglected consideration of the functional overtones in our apprehension of motion. The variables of time and intensity of exposure, interexposure interval, and spatial separation do not by any means exhaust the variance to be found in the movement experience.

REFERENCE

Dixon, W. J. & Mood, A. M. (1946). The statistical sign test. *J. Amer. Statist. Ass.* XLI, 557–66.

Ellis, W. D. (1938). *Source Book of Gestalt Psychology:* (*a*) pp. 161–72, (*b*) pp. 149–59, (*c*) p. 154, (*d*) pp. 149–160.

Köhler, W. & Wallach, H. (1944). Figural after-effects. *Proc. Amer. Phil. Soc.* LXXXVIII, 269.

Neff, W. S. (1936). A critical investigation of the visual apprehension of movement. *Amer. J. Psychol.* XLVIII, 1–42.

Orlanski, J. (1940). The effect of similarity and difference in form on apparent visual movement. *Arch. Psychol., N.Y.*, p. 246.

Stevens, S. S. (ed.) (1951). *Handbook of Experimental Psychology.* New York: John Wiley; London: Chapman & Hall.

Willey, C. F. (1936). Directional variations of apparent movement. *J. Exp. Psychol.* XXIX, 655–73.

(*Manuscript received* 23 *July* 1953)

THE CHANGE OF MOVEMENT TIMING WITH AGE

By W. T. SINGLETON

Nuffield Research Unit into Problems of Ageing, The Psychological Laboratory, University of Cambridge

I. *Introduction* (pp. 166–167). II. *The four-choice experiment* (pp. 168–170). III. *The two-choice experiment* (pp. 170–171). IV. *Discussion* (pp. 171–172). *References* (p. 172).

I. Introduction

There appears to be a widespread belief that older people are slower than younger ones in the speed at which they execute movements, rather than in the speed at which they make decisions and initiate actions. The evidence for this opinion originates from the work of Miles (1931). Using hundreds of subjects on a variety of tests, he was able to show convincingly that, although decrements with age were always surprisingly small, they were more apparent in what he called 'motility measures' than in 'reaction tests'. He defined a motility measure as one including movement only, thus distinguishing it from a reaction measure, which 'included also the strictly cortical elements of comprehension, selection, and readiness for response plus a brief movement' (*op. cit.* p. 632). For reaction tests he used the standard reaction time procedures for both hand and foot, for motility measures he used handle-winding, and a test in which the subject had to move his hand from a key, take a pencil from one vertical hole, insert it in another and return the hand to the key. In recent years, results have accumulated which are not readily reconcilable with this classical view.

Szafran (1951) tested ten subjects in each age decade from twenties to fifties, on an aiming task. The subject sat in the centre of a wooden framework and moved a stylus from a central resting point to any one of a number of 14 in. targets, with 1 in. metal centres, distributed on three sides of him. The task was given under two conditions—with and without vision. When subjects could use direct vision, the time taken to move from the resting point to contact with the target increased slightly with age, but not significantly, and comparatively less than the various stationary components, such as time to initiate response. In this condition the subjects all moved directly to the centre of the target. Without the aid of vision, subjects tended to move the stylus generally on to the particular target and then spent a comparatively long period on target searching for the target centre. In this condition, the time to move from the resting position to the target showed a consistent though not significant decrease with age while other time components of the task increased considerably with age.

Leonard (1952) used a task in which the Display consisted of five lights arranged in a regular pentagon, and the Control was a similar arrangement of metal disks inset in a wooden panel with a sixth disk at the centre of the pentagon formed by the outer disks. The subject moved a stylus from the centre disk to the outer one corresponding to the particular light on the Display. Only one light was on at any time. When the stylus was returned to the centre after touching the indicated disk, the old stimulus light was extinguished and another one appeared. Thus, the task was a serial one involving sliding the stylus to and fro between the centre disk and the outside ones, following unpredictable signals given on the Display. Using a group of young subjects (average age about 25), and a group of older subjects (average age about 65), he found that while all the time components of the task increased with age, there were indications that the speed decrement was greater for the times on the centre disk and on the outer disks, than it was for the transit times between disks. The disks were 1¼ in. in diameter, and the distance between the edge of the centre disk and the edge of any one of the outer disks was 1¾ in.

Both Leonard and Szafran were of the opinion that, although the results in the above experiments were not conclusive, they were consistent with the hypothesis that older people's performance differs from that of younger ones more in the time taken to make decisions than in the time taken to carry out actions.

The findings in all the experiments so far described are easier to reconcile if it is remembered that Miles's distinction between 'motility' and 'reaction' is very difficult to realize in a practical experimental situation. There is both psychological and physiological evidence to support the view that when a manual skill is being performed, a continuous stream of information is passing from the peripheral to the central parts of the nervous system. The concepts of a discrete stimulus—a physical change in the environment which produces a new pattern of sensation which then enters the central mechanisms, and, after a reaction time interval, is translated into muscular action—is a gross over-simplification. The time of an overt movement is very rarely a function of the relevant peripheral processes alone. A movement involves a continuous inflow of nerve impulses both through the proprioceptive system and usually through the optic system, since the eyes are focused on either the moving member or the destination or both. If the movement has some unfamiliar aspects, then the time it takes depends on the speed at which information can be received and further 'orders' retransmitted through the effector system. In other words, a motility measure is usually just as much an index of time taken by central processes as is a reaction measure.

The central processes acting in relation to a movement are essentially concerned with guiding the moving limb, that is, controlling its direction and velocity at any time. If a certain degree of guiding is imposed externally, then it is reasonable to postulate that there will be a corresponding decrease in the extent of central involvement during the movement. One way in which this can be achieved is to ask the subject to move a lever for which the possible directions and amplitudes of movements are mechanically fixed, rather than simply presenting him with starting and ending points and allowing him to find his own path between them.

The author measured the time for an 18 in. movement under two conditions:

(*a*) The subject had to move a stylus to and fro between two disks, each 1 in. in diameter. He was to move as quickly as possible for about half a minute and to take care not to miss touching a disk. The average time of a group of subjects for a single cycle of movement was about 1 sec.

(*b*) The subject grasped a knob at the end of a lever. The lever was constructed so that the knob could be moved only 18 in. in a straight line, with sprung stops at each end of the traverse. The subject was told to move as quickly as possible from one end to the other for about half a minute. The average time of a group of subjects for a single cycle of movement was about half a second.

Detailed results of this experiment, including changes with age, will be published later.

Here it is mentioned as confirmatory evidence that, when a movement is partly guided by external physical constraints, the maximum speed is very much higher than when more control has to be exercised by the subject.

The experiments already cited can now be reconsidered in the light of the following propositions:

(*a*) That the central involvement in action decreases as the accuracy of action adequate for the task decreases.

(*b*) That differences in performance between older and younger people become more marked as the central control required during the performance increases.

All the tasks used by Miles as motility measures were such that some changes of direction of movement were included in the time intervals measured. It may be that it was this characteristic which produced the decrement with age which he found. In the Szafran experiments, there was some decrement with age when the movement was directed towards the small target centre, but no decrement when conditions were altered so that the subject could only hope to move generally on to the large target area. In the Leonard experiment, the relative dimensions of the targets and their distance apart were such that only a very low degree of movement accuracy was sufficient to achieve the desired result. In this case the change with age in movement time between disks was small compared with that found in time spent on the disks where changes in the direction of movement were necessary.

To obtain further evidence in support of this point of view, the following experiments were carried out. The apparatus was developed from that first designed by Leonard (1952); the data reported were obtained from details of performance recorded on an electronic-digital machine—the S.E.T.A.R. (Welford, 1952).

II. THE FOUR-CHOICE EXPERIMENT

Apparatus. The Display consisted of four lights arranged in a square, with one diagonal vertical. The Control was a lever, which could move along four directions from a centre point as shown by the shaded area in Fig. 1. Each light signal indicated a particular direction of movement. During a trial, a stimulus light could be extinguished by moving the lever to the appropriate E; when the lever was then returned to C a new stimulus light would appear. The Control and Display were alined so that a light could only be put out by moving in the direction opposite to that where the light was on, e.g. a light on at the left meant move to the right, a light on at the top meant move towards the body and so on. In this way the task was made more difficult than it would have been had a directly alined Display and Control been used (Singleton, 1953).

For the purpose of this paper it is sufficient to note that the task was to move a lever C to E, E to C, C to E, and so on in response to light signals. The lever was sprung to facilitate stopping and changing direction at C. A movement C to E required a knob displacement of about 4 in.

Method. Ten male subjects in each age decade were used. They were selected at random from acquaintances of the experimenter, visitors to the Laboratory, University employees, etc., with the one criterion that in each age decade there should be five subjects of approximately University or professional status, and five at about the average educational level of the wage-earning community. Thus in the twenties five subjects were students and five were naval ratings. Each subject was given six trials on the task; a trial consisted of sixty-four correct responses.

Results. For each correct response three time measures were recorded: the time spent at *C*, that is the time between the appearance of the stimulus and the initiation of the response; the time spent in moving from *C* to *E*; and the time spent between arriving at *E* and arriving back at *C*. When a subject made a mistake (that is, started from *C* towards an *E* which would not extinguish the light), his centre time for that response was omitted. The distribution of errors with age is not easy to analyse, but there is certainly no marked linear trend of either increase or decrease from 'twenties' to 'sixties'.

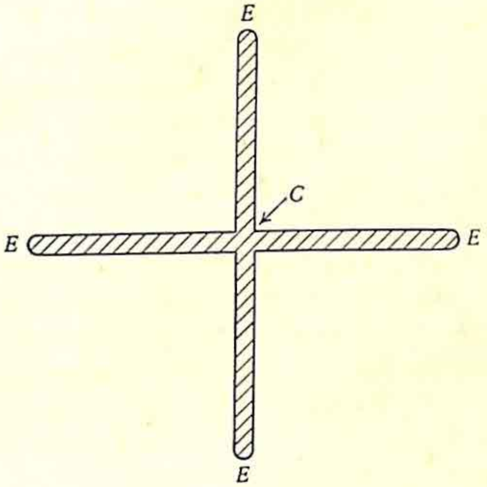


Fig. 1. The shaded area represents the possible directions of movement in the Four-Choice Experiment.

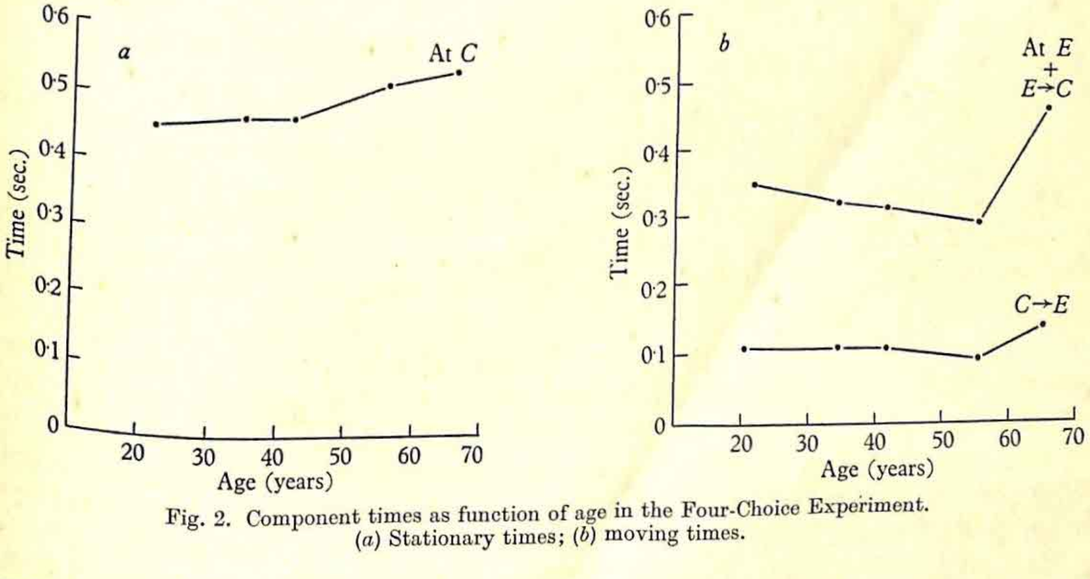


Fig. 2. Component times as function of age in the Four-Choice Experiment.
(*a*) Stationary times; (*b*) moving times.

All three measures taken during the sixth trial are shown as means for subjects in each age decade in Fig. 2*a*, *b*. The time (at *C*) increases linearly with age, the time (*C* to *E*) remains reasonably constant with age. The time (at *E* plus *E* to *C*) appears to decrease from 'twenties' to 'fifties', and shows an abrupt increase for 'sixties'. This unusual function is the subject of further experiments not immediately relevant to the present issue, which will be described in a later paper. It may be noted, however, that it is consistent with the results of Szafran (1951), who found a small but regular decrease in movement time from 'twenties' to 'fifties' and with Leonard (1952), who found an increase in movement time when a group of 'twenties' was compared with a group of 'sixties'.

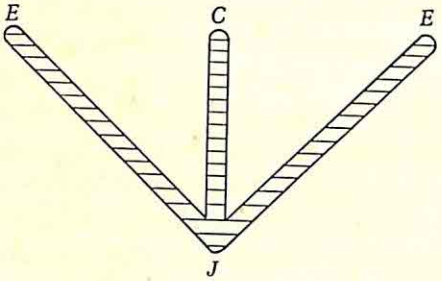


Fig. 3. The shaded area represents the possible directions of movement in the Two-Choice Experiment.

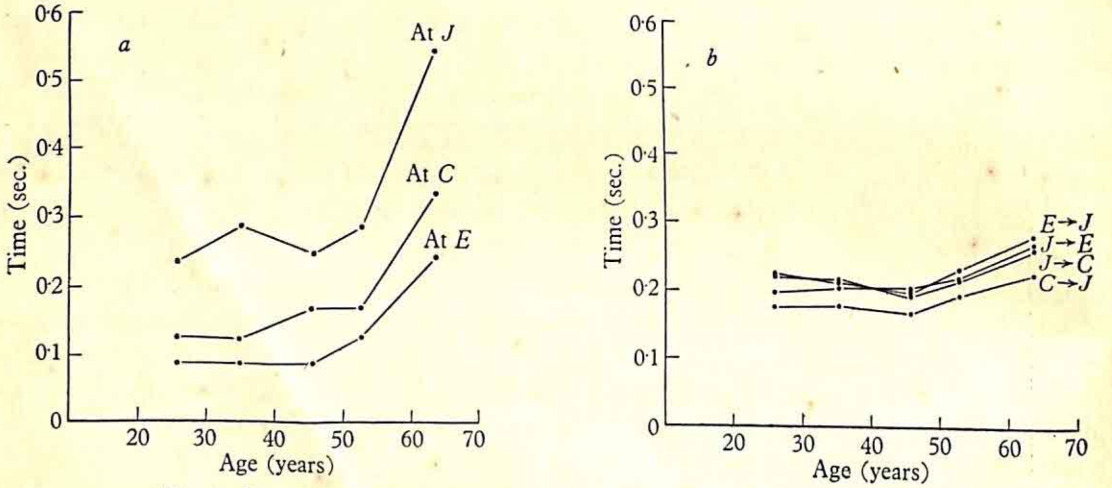


Fig. 4. Component times as functions of age in the Two-Choice Experiment.
(*a*) Stationary times; (*b*) moving times.

III. THE TWO-CHOICE EXPERIMENT

Apparatus. The Display consisted of two lights arranged horizontally about 9 in. apart. The Control was a lever which could move along the directions shown by the shaded area in Fig. 3. During a trial the stimulus light could be extinguished by moving the lever to the appropriate *E*, when the lever was returned to *C* via *J* a new stimulus light would appear. The Display and Control were directly alined so that, for example, a light on the left meant move from *C* to the left *E*. The lever was lightly sprung in the general direction *JC*. A movement *C* to *J* required a knob displacement of about 4 in., from *J* to *E* about 6 in.

Method. Forty-eight male subjects were used, fairly evenly distributed over the age range from 20 to 70 years. All were employees of a Lancashire cotton mill. Each subject was given one unscored practice trial followed by two experimental trials on the task. A trial consisted of forty-eight correct responses.

Results. For each correct response eight time measures were recorded. The time spent at *C*, moving from *C* to *J*, at *J*, moving from *J* to *E*, at *E*, moving from *E* to *J*, at *J*, and moving from *J* to *C*. Almost all the errors made were manipulatory ones at *J*; such error times at *J* were omitted in calculating the means. There was no marked error trend with age. The data of Fig. 4 were extracted from records of the final trial.

The changes with age in the four moving times are all of the same form. The stationary times against age curves are all three of the same form but, as a class, appear to be different from the moving times. The 'at *J*' component is the sum of '*J* going out' and '*J* returning'. These latter do not appear to differ from each other as functions of age. It will be noted that the curve 'at *E* plus *E* to *C*' of Fig. 2 is of the same form as the stationary times in Fig. 4, which suggests that the 'at *E*' component is controlling this function at least at the upper end of the age scale. The stationary time 'at *C*' in Fig. 2 is not the same shape as the stationary times in Fig. 4, presumably because the former is determined by time necessary to receive new information rather than time to change direction. This may be one case of a phenomenon which merits further investigation. The central processes associated with changing a movement direction appear to overlap in time with perceptual processes determining the next direction of movement.

IV. DISCUSSION

The manipulative skills performed in ordinary life are usually such that information concerning what must be done next becomes available while other movement is being executed, and some time before it is necessary to take appropriate action. This introduces an ambiguity in the results obtained by recording the time components of such performance. It is not known whether a particular time interval is a function of the 'tactical' action of moving a limb from one point to another, or of some more 'strategic' process concerned with the planning of future action. The tasks described above were essentially artificial, in that information concerning the next action only became available when the previous action was complete, and the degrees of freedom in carrying out movements were quite strictly determined by the apparatus.

Within these limitations, the results of the two experiments suggest that the performance of older men on simple, but somewhat unfamiliar perceptual motor skills, tends to be slower than that of younger men. This difference in overall speed is not due to slower movements so much as to longer times spent at points where the movement direction must be altered.

It seems reasonable to make a more general statement in the form of the following hypothesis. If the activity is such as to impose no undue stress, the difference between the performance of older and younger groups of people is due to extra time which the former use to absorb information at a central level rather than to time taken up directly because of some deterioration of the peripheral receptor and effector processes.

The phenomena discussed in this paper may be related to several others which have been revealed by the improved recording techniques available in recent years. Rubin, von Tebra & Smith (1952) have reported that '...it is mainly the manipulation component of a motion pattern that is influenced by practice and not the travel component'. Seymour (1954) has suggested that '...increased speed of performance in accomplishing such industrial tasks (simple manual skills) results not so much from faster movement as from a briefer duration of stationary therbligs or elements'. In relation to certain manipulative skills, Singleton (1953) wrote that 'the important determinant of performance and indicator of fatigue is not the time required to execute movements, but the time between movements'. All the experiments reported or referred to in this paper indicate the importance of separating into their various components the total times of perceptual-motor tasks. In particular, they point to the crucial role which the time occupied by central mechanism plays in the performance of such skills.

I am indebted to Mr A. T. Welford for advice on the work described in this paper. I should also like to thank Mr L. V. Green of the Dunlop Rubber Company and the staff of the Dunlop Cotton Mills, who arranged for and acted as subjects in the Two-choice experiment.

REFERENCES

LEONARD, J. A. (1952). Some experiments on the temporal relation between information and action. Unpublished Ph.D. Thesis, Cambridge University.

MILES, W. R. (1931). Measures of certain human abilities throughout the life span. *Proc. Nat. Acad. Sci., Wash.*, XVII, 627–33.

RUBIN, R., VON TEBRA, P. & SMITH, K. V. (1952). Dimensional analysis of motion. III. Complexity of movement of pattern. *J. Appl. Psychol.* XXXVI, 272–6.

SEYMOUR, W. D. (1954). Manual skills and industrial productivity. *J. Instn Prod. Engrs.*

SINGLETON, W. T. (1953). Deterioration of performance on a perceptual motor task. Pp. 163–72 in W. F. Floyd and A. T. Welford (ed.) *Symposium on Fatigue*. London: H. K. Lewis and Co. Ltd.

SZAFRAN, J. (1951). Changes with age and with exclusion of vision in performance at an aiming task. *Quart. J. Exp. Psychol.* III, 111–18.

WELFORD, N. T. (1952). An electronic digital recording machine the S.E.T.A.R. *J. Sci. Instrum.* XXIX, 1–4.

(*Manuscript received* 9 *February* 1954)

COGNITIVE CHANGES IN THE FEEBLEMINDED[1]

By A. D. B. CLARKE and A. M. CLARKE
Psychology Department, The Manor Hospital, Epsom

I. *Introduction* (p. 173). II. *Frequency of* I.Q. *changes* (pp. 173–175). III. *Hypotheses to explain results* (p. 175). IV. *The hypothesis of adverse environmental influence* (pp. 175–177). V. *Conclusions* (pp. 177–179). *References* (p. 179).

INTRODUCTION

An earlier paper by the writers (1953) has discussed the general question of I.Q. variability, and evidence was presented indicating that during the period of mental growth I.Q. constancy for the individual over long periods of time is the exception rather than the rule. The thesis was advanced that the I.Q. must be regarded as an objective and relatively reliable estimate of a person's present intellectual status, but that for predictive purposes, even though better than subjective judgement, the results of intelligence tests will be unreliable for a considerable proportion of individuals. Thus, Honzik, MacFarlane & Allen (1948) found that only 15% of a group of 153 children changed in I.Q. by 9 points or less between the age of 6 and 18, while 58% changed 15 or more I.Q. points.

The present paper examines in detail an investigation of I.Q. changes in the certified feebleminded. It began with the observation that among a group of eighteen adolescent and young adult patients there had occurred in some cases quite large increases in I.Q. on the Wechsler Test, after a time interval of about 18 months. These changes ranged from a decrease of 5 to an increase of 25 points upon retest, with a mean increase of 9·3, S.D. 7·3. A control group of eighteen patients, all new admissions, was matched with the first group for initial I.Q. and age, and retested after a time-interval of only 3 months, in order to determine the maximal effect of test practice, errors of measurement and under-estimation at the first test. It was found that an average increment of only 4 points occurred under these conditions, this being significantly lower than the increase shown by the first group.

It was further noted that although researches existed in which I.Q. changes in mental defectives had been studied, rarely had such workers as Burt (1947), McKay (1942), Roberts (1945), Spaulding (1946) and Guertin (1949) done anything more than to demonstrate that such changes occur, and in some cases to suggest hypotheses to explain irregularity of intellectual growth. In most cases these hypotheses seem never to have been submitted to experimental verification.

II. FREQUENCY OF I.Q. CHANGES

Since it was known that the first small group was not a representative sample of the feebleminded hospital population, it was decided to plan a research using a larger number of patients to determine how frequently I.Q. changes could be expected to occur among the younger certified feebleminded, and, more important, why such changes take place.

As a preliminary, a note was made of every patient who had been first tested with the

---

[1] This article is based upon papers read to the Mental Deficiency Section, The Royal Medico-Psychological Association, 5 November 1952, and to the British Psychological Society, 13 April 1953.

Wechsler about 2 years previously, and a detailed examination of reasons for referral showed that the fifty-nine patients still available were a representative sample of the younger feebleminded hospital population. These individuals were then retested, and Table 1 shows the results. It will be noted that the increase was significant beyond the 0·1% level of confidence. It should be further noted that the increments were general throughout all the subtests of the Wechsler Scale, apart from Digit Span, where the increase did not reach the 5% level of significance. It had been predicted that there would be a greater increase on the performance items than on the verbal subtests. The mean verbal weighted score increase was 3·203, S.D. 4·405, while that for the performance subtests was 4·729, S.D. 6·161. The difference between these means is, however, only just significant at the 5% level, using the one-tail test.

Table 1. *Results of test and retest on Wechsler, Form I, for fifty-nine patients retested after an average interval of* 27 *months*

	Original full scale I.Q.	Retest full scale I.Q.	Difference
Mean	66·2	72·7	6·5
S.D.	14·0	13·4	6·2
N	59	59	59

t (correlated means) = 8·074, significant beyond the 0·1% level

I.Q. range at original test: 35 to 98
I.Q. range at retest: 40 to 97
Age range at retest: 14 to 50 years, mean 23·5, S.D. 8·1*
Period between test and retest: mean 27 months, S.D. 6 months.
Range of I.Q. changes: −7 to +25 I.Q. points.

* Only nine patients were over 30 years of age.

It was felt necessary to augment the control group, and therefore a further eleven patients were retested after a short time-interval; for this larger control group of twenty-nine patients the mean increase was 4·1, S.D. 3·8 I.Q. points, this increase being significant at the 0·1% level. A comparison of the distribution of changes in the main group and the control group will be found in Table 2.

Table 2. *Comparison of* I.Q. *changes for the main group and control group*

Change	Main group		Control group	
− 7 to 0 I.Q. points	9/59 = 15%		4/29 = 14%	
+ 1 to +7 I.Q. points	23/59 = 39%	100%	19/29 = 65%	100%
+ 8 I.Q. points or more	27/59 = 46%		6/29 = 21%	
+10 I.Q. points or more	17/59 = 29%		1/29 = 3%	
+15 I.Q. points or more	7/59 = 12%		0	
+25 I.Q. points	1/59 = 2%		0	

Inspection of the Table shows clearly the greater gains made by the members of the main group. The first question raised at the outset can now be answered; almost half of the patients restested after 2 years showed gains of 8 points or more, this figure having been chosen as the base-line for significant cognitive improvement. It is obvious, moreover, that for the majority of such patients the increments occurring over this time-interval do not represent the total change which the individual has already made or will make in the future. Honzik *et al.* (1948), for example, have pointed out that some subjects show consistent trends in increment or decrement over many years. The results thus give a picture of a considerable proportion of the certified feebleminded advancing towards intellectual normality at ages when mental growth is commonly assumed to have ceased.

III. HYPOTHESES TO EXPLAIN RESULTS

The second important research problem was to determine the reason for these increments, and the following main hypotheses were investigated:

(1) That the results can be explained entirely in terms of test practice, errors of measurement and underestimation at the time of the first test. The results from the control group of twenty-nine patients retested after an average of 3 months, who showed only a 4-point increment on the average, dispose of this hypothesis.

(2) That those individuals who made the big increases would be those who were new admissions to the hospital at the time of the first test and were underestimated since admission may represent a traumatic experience for the individual. It was found that the increments of those first tested as new admissions were no greater than for those who had been in the institution for a considerable time before the first test. Furthermore, the control group of patients retested after an average of 3 months were mostly new admissions examined within a few days of arrival.

(3) That those who increased would be the younger members of the sample. It was found that the correlation between age and weighted score increment was $-0{\cdot}155$, which is in the expected direction but is not significant.

(4) That there would be a relationship between initial I.Q. and subsequent change. The correlation between full-scale I.Q. on the first test and subsequent change is $-0{\cdot}311$, significant at the 2% level of confidence. There is, therefore, a small but significant tendency for those with lower initial I.Q.'s to make larger increases than those with higher initial I.Q.'s.

(5) The final hypothesis, which had emerged from inspection of the data concerning the first group of eighteen cases, was that those whose condition, irrespective of possible genetic causes, seemed to be related to early adverse environment, would be those in whom the larger I.Q. increases occurred.

IV. THE HYPOTHESIS OF ADVERSE ENVIRONMENTAL INFLUENCE

The first method used to investigate this final hypothesis involved an independent analysis of all documents relating to each patient, and an attempt was made to determine the aetiology of the deficiency in each case. The categories included disease, adverse environment, heredity and various combinations of these. The I.Q. increments were then plotted for the members of each aetiological group and all were insignificantly different from one another, apart from the disease group for which the mean increment, as might be expected, was only half an I.Q. point. There were, however, only four patients assigned to this group, a number, of course, too small to warrant general conclusions. The failure of aetiological classifications to correlate with cognitive changes could be the effect of either of the following factors: first, that, as Penrose (1949) points out, it is extremely difficult to determine the precise aetiology of mental deficiency, except in the relatively few clear-cut cases of disease and low-grade mental defect. The problem of aetiological classification becomes apparent when one remembers that while the incidence of severe defect is equal throughout society, the feebleminded, who form the vast majority of mental defectives, are to be found mainly in the socially deprived groups in which it might be argued that both adverse heredity and adverse environment may be the factors responsible. Secondly, that, in fact, cognitive improvement can be expected in all feebleminded patients irrespective of aetiology, with the probable exception of those whose condition is attributed to disease or injury.

On the basis of previous work the latter of these two possibilities seemed to be unlikely and therefore further work was undertaken to elucidate the failure of aetiological classification to be of significance. It was argued that there must be factors related to these I.Q. changes, either in the past or the present histories of the patients concerned. There appeared to be nothing relevant in the present history, so once again all records of past history were examined, and all specific factors which could possibly be relevant were separately tabulated: e.g. number of sibs, illegitimacy, parental separation, parents said to be of poor mentality, early separation from the parents, institutional upbringing, N.S.P.C.C. reports, absence from school, illness, etc. Inspection of the results seemed to show that certain indications of an early bad home, irrespective of possible genetic factors, corresponded with the amount of I.Q. change. An attempt was therefore made to formulate valid criteria of early adverse environment. These are given in rough order of importance:

(1) N.S.P.C.C. or Crusade of Rescue adverse reports or intervention, or statement in the case notes 'Family known to the N.S.P.C.C.'.
(2) Parental attitude to the child said to be antagonistic or uninterested. Parental cruelty and fear of the parents by the child.
(3) No fixed abode. Living in shelters with the parents or on the street.
(4) Orders appointing someone as 'Fit Person' to take charge of the child. This normally means the removal of the child from the parent by some authority.
(5) Statement in the report that 'Home conditions are bad' or 'very bad'.
(6) Statement in the report that the child has been considerably neglected.
(7) Irregular attendance at school due to parental neglect or lack of interest.
(8) Home said to be dirty and neglected.
(9) Gross poverty.
(10) History of crime in parents or sibs, gross intemperance, etc.
(11) Dietary deficiency: rickets.
(12) Child found begging or wandering or missing from home.

The presence of two or more of these criteria was taken as evidence of a very bad home, and in most cases thus classified several of these somewhat overlapping criteria were present. This may be illustrated by the bare outlines of two typical case histories:

(*a*) Amy and her twin sister, now aged 18, were admitted to a hospital with measles and a wasting disease at the age of 18 months. At the age of 3 Amy was placed in a convent as a resident owing to the poverty of the family. She was removed at the age of 12 against the advice of the Crusade of Rescue. The mother ill-treated both twins badly, and the school reports state that they suffered insidious torment at home; they were sent to school in a shocking condition, and when the teachers supplied clothes to replace the children's rags, these clothes had to be kept at school or the mother would remove them. Amy's progress at school was very adversely affected by these circumstances, and the child herself was pitiful in her requests to be taken away from home.

(*b*) Stanley, now aged 21, was said to have had 'fits' between 1935 and 1940, but there are no details of them. He suffered from rickets as a child; school reports state that his educational attainments were very poor, but he was obedient and quiet. He wandered, and school attendance was very bad. At the age of 12 he was brought before the Juvenile Court as being in need of care and protection. He was said to be terrified of his father, and was never allowed to have enough sleep. The parents were that year prosecuted for his non-attendance at school. The father was stated to be of low mentality, the home to be dirty and neglected, and the mother to be illiterate, callous and intemperate. The parents separated shortly thereafter.

When these twelve criteria had been formulated, they were applied to all fifty-nine case histories, and it was found that those with very bad homes were mainly those who had made the larger I.Q. increments. In order to check upon this finding, and to exclude any possibility of subjective bias, it was necessary to call in an outside investigator to carry out the ratings independently, and this Dr J. Tizard of the Unit for Research in Occupational Adaptation, Medical Research Council, very kindly consented to do. Having no knowledge whatever of either the patients or the I.Q. changes which they had made, Dr Tizard analysed the contents of all the fifty-nine case histories with the twelve criteria in mind, and assessed presence or absence of very bad homes. During the 3 days when this task was in progress, there was no discussion whatever of the cases or of the research between Dr Tizard and the writers.

Upon completion it was found that there was agreement on fifty-five out of the fifty-nine cases; in the four instances where there was disagreement, Dr Tizard's judgement was left unaltered and accepted as final. There were thus two groups of patients, those whom the external investigator considered to have very bad homes and those against whom there was either no evidence, or the evidence was inconclusive, or there existed evidence pointing to good homes. Details of the I.Q. increments for the members of the two groups are shown in Table 3, together with control group data.

Table 3. *Changes in* I.Q. *points*

	Very bad homes	The remainder	Control group
Mean change	+9·7	+4·1	+4·1
S.D.	6·3	4·9	3·8
N	25	34	29

t (uncorrelated means) = 3·864, significant beyond the 0·1 % level.

Distribution of changes

Change	Very bad homes		The remainder	
− 7 to 0 I.Q. points	1/25 = 4 %	100 %	8/34 = 24 %	100 %
+ 1 to +7 I.Q. points	6/25 = 24 %		17/34 = 50 %	
+ 8 I.Q. points or more	18/25 = 72 %		9/34 = 26 %	
+10 I.Q. points or more	12/25 = 48 %		5/34 = 15 %	
+15 I.Q. points or more	7/25 = 28 %		0	
+25 I.Q. points	1/25 = 4 %		0	

The difference between the gains made by the two groups is significant at the 0·1 % level. It will be noted further that the group of thirty-four patients who were not classified as having had very bad homes made the same sort of gains as the members of the control group; in other words, their increments were of the order one would expect from the maximal effects of practice, errors of measurement and possible underestimation at the time of the first test.

V. CONCLUSIONS

These results show clearly two things: first, the very close agreement, 93 %, between independent ratings of bad home points to the objectivity of such assessments; and secondly, that these criteria are effective in discriminating on the basis of early history between those who later in life had I.Q. increments, and those who showed no such large changes.

It should be remembered that the majority of those who are certified as feebleminded come from what the man-in-the-street would consider to be bad homes. In the present sample of fifty-nine individuals, for instance, in only six cases was there evidence of good homes. What the present investigation has shown, therefore, is that in a population in general drawn from bad homes those with the worst are distinguished from those with the less bad homes by their later cognitive improvement. It is consistent with this main finding to suggest that the environment which is really antagonistic towards the child retards mental development for many years. Later, however, this retardation begins to fade, I.Q. increments thus occurring. There is as yet no certain evidence to show whether this is the effect of delayed maturation or of the relatively better and more stable environment of the institution, or a combination of both. In view of the fact that length of time in the institution is not associated with the increases, it seems likely that these changes are more the effect of removal from a very adverse environment than of entry into a relatively better one.

Retrospective validation such as has been described carries with it the possibility of error by capitalizing upon chance relations existing within the records of those investigated. It was necessary, therefore, to find out whether, by using the twelve criteria of very bad homes, it was as possible to predict I.Q. increments as it had been to discriminate retrospectively. From the twenty patients who had been first tested during the first few months of 1951 and who had not yet been retested as members of the control group, eleven were selected, consisting of those in whose records there were clear indications of very bad homes and those in whom this was contra-indicated. Before any retesting was carried out, definite predictions were made that there either would be or would not be a significant increment. Six patients for whom increments had been predicted showed increases on retest of: +14, +9, +10, +10, +4 and +5, with a mean of 8·7 I.Q. points, and the five in whom increments had been contra-indicated showed the following changes: −5, −2, +7, −1, +4, with a mean of 0·6 of an I.Q. point. On this very small sample, therefore, prediction based on the criteria is fairly accurate. It is, of course, intended later to check this finding independently on a larger sample.

The findings can now be summarized and the implications considered. The results must be seen against the background of the several hundred researches which indicate I.Q. variability as a normal feature over considerable lengths of time. In this investigation, however, the changes which occurred were mostly in the upward direction and were often large in relation to the relatively short time-interval of about 2 years between test and retest, almost half of the fifty-nine subjects having increased by 8 points or more, the maximum individual increment being 25 points.

While many previous researches have suggested hypotheses to account for I.Q. variability, seldom have such hypotheses been carried to the stage of experimental verification. An attempt was made in the present investigation to remedy this defect. It was shown that I.Q. increments are associated with early very adverse environment, and that intellectual subnormality among such deprived people as the certified feebleminded is not necessarily a permanent and irreversible condition. Obviously in many cases the increments in I.Q. over a period of 2 years do not represent the total change for the individual.

The main question which this research has raised is as follows: having shown that I.Q. increments occur in a large proportion of the certified feebleminded, and that these increases are associated with early deprivation, an investigation must now be planned to find out how such relatively spontaneous changes can be both accelerated and improved.

The writers wish to thank Dr J. F. MacMahon, Physician Superintendent, for his very great help and encouragement during the entire course of the research. Dr J. Tizard, M.R.C. Unit for Research in Occupational Adaptation, played a vital part in one aspect of the investigation, and Mr M. B. Shapiro, Senior Lecturer, Psychology Department, Institute of Psychiatry (Maudsley Hospital) made stimulating suggestions and criticisms. Mr A. Ehrenburg, Statistical Laboratory, Maudsley Hospital, gave useful advice, and Mr G. J. F. Singleton kindly performed the statistical computations.

REFERENCES

BURT, C. (1947). *Mental and Scholastic Tests.* London: Staples Press.
CLARKE, A. D. B. & CLARKE, A. M. (1953). *Lancet,* II, 877.
GUERTIN, W. H. (1949). *J. Clin. Psychol.* V, 414.
HONZIK, M. P., MACFARLANE, J. W. & ALLEN, L. (1948). *J. Exp. Educ.* XVII, 309.
MCKAY, B. E. (1942). *Amer. J. Ment. Defic.* XLVI, 496.
PENROSE, L. S. (1949). *The Biology of Mental Defect.* London: Sidgwick and Jackson, Ltd.
ROBERTS, A. D. (1945). *Amer. J. Ment. Defic.* L, 134.
SPAULDING, P. J. (1946). *Amer. J. Ment. Defic.* LI, 35.

(*Manuscript received* 8 *September* 1953)

A THEORY OF HUMAN CURIOSITY

By D. E. BERLYNE

University of Aberdeen

I. *Introduction* (pp. 180–181). II. *The sequence of events* (pp. 182–184). III. *The role of conflict* (pp. 184–189). IV. *Summary* (pp. 189–190). *References* (pp. 190–191).

I. INTRODUCTION

Few phenomena have been the subject of more protracted discussion than human knowledge. Yet this discussion has usually paid little attention to the motivation underlying the quest for knowledge, with the result that two important questions still confront us. The first question is why human beings devote so much time and effort to the acquisition of knowledge. Sometimes there is some obvious drive to whose satisfaction it can contribute. But, strangely enough, many of the queries that inspire the most persistent searches for answers and the greatest distress when answers are not forthcoming are of no manifest practical value or urgency. One has only to consider some of the ontological inquiries of metaphysicians or the frenzy of crossword enthusiasts to be convinced of this. The second question, and the main concern of the present article, is why, out of the infinite range of knowable items in the universe, certain pieces of knowledge are more ardently sought and more readily retained than others.

Modern learning theory leads us to look for motivational variables to answer these questions, and a drive which is reduced by the reception and subsequent rehearsal of knowledge is what we generally call 'curiosity'. However, we must draw a distinction between this curiosity and the curiosity drive that has been studied in lower animals (Berlyne, 1950). In the case of the rat, for example, there appears to be a drive which is aroused by novel stimuli and reduced by continued exposure to these stimuli. Its reduction reinforces exploratory activity, i.e. activity, such as approaching and examining the stimulus-objects, which increases stimulation of the animal's receptors by them. Now, similar exploration is undoubtedly elicited by strange objects in adult and especially infant human beings. But in an animal as well endowed for learning and remembering as the human adult, exploration is bound to leave a stock of permanent traces in the form of symbolic representations ('pure stimulus acts' or 'cue-producing responses'), which are manifestations of what we call 'knowledge'.

The curiosity which leads to increased perception of stimuli and the curiosity whose main fruits are knowledge may well turn out to be closely related. But, as we are using different defining operations for them, we shall have provisionally to use two different terms. We shall therefore call the first '*perceptual curiosity*' and the second, which is our concern in this article, '*epistemic curiosity*'.

An account in behaviour-theory terms of the nature and origins of knowledge has been offered elsewhere (Berlyne, 1954). According to this account, which draws on concepts introduced by Morris (1946) and by Osgood (1952), knowledge consists of *habits mediating believed, designative symbols*, which form sequences (trains of thought). The actual course followed by a train of thought is determined jointly by (1) *cue-stimuli*, which include external stimuli (S) and the self-stimulation (s) resulting from previous items in the sequence and (2) *motivational stimuli*, which include drive-stimuli (S_D) and the goal-stimuli (s_G) produced by fractional anticipatory goal-responses (r_G).

The learning that produces knowledge can clearly be biologically helpful because (1) it can enable goal-directed behaviour to be more efficient through being better prepared for what is impending, and (2) it can enable warning signals to be recognized, so that danger can be avoided (Mowrer, 1939). But these effects are usually delayed, so that their advantages do not explain what reinforces the learning in question. For long-term consequences to influence behaviour, it is necessary for symbols to be used in such a way that the reinforcements of fear-reduction and secondary rewards can be brought to bear (Mowrer & Ullman, 1945). The drives that are reduced by knowledge are thus largely the 'coexistent emotional components', which, as Ullman (1951) argues, we must assume to be present in all primary drives and to be capable of functioning anticipatorily. These components are what are called 'fears'—Ullman speaks of 'shock-fear' and 'hunger-fear', for example—and knowledge can lead to fear-reduction in various ways: (1) by depicting the future situation as a desirable one (reassurance), (2) by reducing 'fear of fear' or 'fear from a sense of helplessness' (preparation) (Mowrer & Viek, 1948), or even when the outlook is hopeless, (3) by reducing the increment of drive (conflict) due to uncertainty (the comfort of 'knowing the worst'), or (4) extinction of fear by repeated exposure to frightening stimuli ('getting used' to unpleasant prospects).

But our main concern is with the second question, viz. the question of the factors underlying the selectivity of epistemic curiosity. Why does an individual seek or learn one piece of knowledge rather than another? Representatives of various schools of psychology have provided hints but scarcely more:

(1) *Psychoanalysis*. The writings of Freud (e.g. 1905) and his followers (e.g. Abraham, 1927) make it clear that psychoanalysts would attribute the desire to know to any of several 'component drives' of the libido—scoptophilic, oral-incorporative, oral-sadistic, anal-aggressive or anal-retentive—according to the direction it takes. But this leaves many questions unanswered. How are we to predict when one of those 'component drives' will find an outlet in curiosity, and how intensely? And which particular items of knowledge will be sought?

(2) *Gestalt psychology*. Although the Gestalt psychologists have not produced a systematic account of curiosity, it is not difficult to guess how such an account would go. They explain much of behaviour by the '*principle of closure*', the tendency to act in such a way as to close a 'gap', whether in a perceived figure or in some other aspect of the 'behavioural world' (Koffka, 1935; Wertheimer, 1945). It is evident that curiosity consists precisely of a drive to fill in such gaps in the subject's experienced representations. But again, we have no definition precise enough to tell us infallibly what will constitute a 'gap', nor which gaps will have precedence over others.

(3) *Reinforcement theory*. The tendency to acquire the verbal or other responses which constitute knowledge is a product of learning, culturally conditioned, according to such reinforcement-theorists as have considered the problem. Dollard & Miller (1950), mention learned drives to 'make a correct report of the environment' and to 'have an explanation' and the punishment that social training, as well as the demands of reality, imposes on those who fail to do so. Skinner (1947) similarly describes how a child learns to emit 'tacts' (i.e. verbal responses controlled by properties of objects or situations)

under the influence of 'generalized reinforcers', particularly approval. Mowrer (1950) appears to identify the acquisition of 'beliefs' (p. 5) and 'knowing that' (p. 268) with the conditioning of emotional responses, but this does not acknowledge the role of symbolic responses in distinguishing pieces of knowledge with similar affective value but different content.

We shall take these treatments as a starting-point, although it is clear that they leave some essential questions unsettled. There is, for one thing, the paradoxical fact that curiosity seems to be evoked most uniformly by situations that are new and strange. They would be the last we should expect to have any influence at all, if it were a matter of generalization from prior training (Berlyne, 1950, p. 71). It will be our contention that conflict supplies the clue to these cases.

II. THE SEQUENCE OF EVENTS

Stimuli which are used to elicit verbal behaviour, unless they resemble the behaviour they call forth or have unique responses, are what Skinner calls '*thematic probes*' (1953, 1947). They can take the form of verbal or non-verbal stimuli, and they can be administered to oneself ('self-probes'), as a reaction to a perceived situation, or come from outside in the form of writing, speech or non-linguistic cues. We can extend Skinner's concept a little, but not, it is hoped, inexcusably, by including under it all stimuli which elicit trains of thought, whether verbal in content or not.

Skinner gives as illustrative cases of thematic probes the stimulus-words of association experiments and the material used in projection tests. But it has been known since the work of the Würzburg school (Ach, 1910) and of Lewin (1917, 1922) that it is not possible to predict what association, if any, will be given to a stimulus without taking into account the 'set', 'determining tendency', 'tension-system', etc., induced by other stimuli, usually instructions. In other words, we must have not only cue-stimuli to act as a starting-point, but also motivational stimuli to limit the responses to the general category required by the task on hand and to supply the persistent motive force for the process. The thematic probe must thus have two parts or aspects with these distinct functions, and the clearest example, as well as probably the commonest in practice, is the *question*. The question of the type called by linguists the 'specific interrogation' (as distinct from the 'yes-or-no question') (Bloch & Trager, 1942) has the two parts easily distinguishable. As an example, we may take the question, 'how does the starfish eat?' We assume that the question, in common with all synonymous questions, evokes mediating 'concepts' or 'meaning' responses (r_m) (Berlyne, 1954). The meanings corresponding to 'starfish eat' act as the cue-stimuli with a patterning effect peculiar to that stimulus-complex; in some cases, such as when the question is put by an authoritative person, they may be tantamount to an assertion that the starfish eats, while in other circumstances the question may be taken to mean 'how, if at all, does the starfish eat?' The group of concepts that act as cue-stimuli we may, following Morris's terminology, call the '*designator*'. On the other hand, the interrogative adverb 'how' produces a meaning which acts as a motivational stimulus. It limits the train of thought to 'how-concepts' and evokes a learned drive-state which motivates the reaction.

When a question is put, whether by the subject himself or by somebody else, and the answer is already known, the appropriate response is made as a reaction conditioned by previous learning to the stimulus-pattern, and this relieves the drive immediately, so that the subject can proceed to some other activity. However, when the answer is not known, the drive will persist, and some sort of trial-and-error process can be expected to follow as with any other drive-state. Of course, the trial-and-error will not be completely random, as it is not even for the rat in the Skinner-box: it will take the form of behaviour resembling what has succeeded in similar situations. The most likely behaviour-sequences to occur are: (*a*) *thinking*—implicit trial-and-error, insightful restructuring (Hull, 1952, chap. 10), stimulus generalization (as in 'deduction' Mowrer, 1950, chap. 11), 'intuition' (Piaget, 1947) and 'magical thinking' (Fenichel, 1945, p. 47); (*b*) *observation*—approach, receptor-adjustment, manipulation of environment, so as to perceive relevant stimuli, culminating in the controlled experimental and other techniques of science; (*c*) *recourse to authority*—asking experts, consulting books or oracles, etc. (Cf. intuitionism, rationalism, empiricism, and authoritarianism (Montague, 1936).)

If the processes lead to a pattern of responses that the subject's prior learning enables him to accept as an adequate answer, then the drive will be reduced. Since drive-reduction follows the rehearsal of the correct answer, the principle of reinforcement (Hull, 1952, postulate IV) implies that the latter will become strengthened as a response to the question. Furthermore, by the reinforcement-gradient principle, it will be learned more strongly than the responses that led up to it, so that in future the question will be followed by the answer immediately, and intermediate steps will be omitted.

If the answer is not arrived at readily by any of the procedures mentioned, then the process may be brought to an end in other ways. Some distraction may occur, i.e. an incompatible response-tendency with a higher reaction-potential may arise, or extinction may supervene. There will be extinction of each line of inquiry as it turns out to be unsuccessful and gives way to another, exactly as in trial-and-error learning (Hull, 1930), and also ultimately the drive-producing responses may be extinguished, so that the subject gives up altogether. It is unlikely that extinction will affect the intervening link, that between the words and the drive, since the motivating power of interrogative adverbs is frequently and partially reinforced in everyday life.

Let us now suppose that the subject fails to hit upon the correct answer in the course of striving for it. And let us suppose that on some future occasion he is told or shown the answer, i.e. exposed to some stimulus-complex which evokes the response he was seeking. We can expect this answer to elicit, by ordinary redintegrative remembering, an internal rehearsal of the question, so that it is *recognized* as the answer he was looking for on the earlier occasions. The stimuli produced by the response of rehearsing the question will thus occur about the same time as the rehearsal of the answer, and the stimuli produced by rehearsal of the answer will be followed closely by reduction of the drive that the question has re-aroused. Thus we can see the answer being learned by reinforcement as a response to the question, so that a new piece of knowledge is acquired.

It will by now be evident that the drive aroused by questions and other thematic probes is, by our definition, a form of epistemic curiosity. And an important consequence follows from the principles of behaviour theory, if our account so far is valid, which gives us a way of measuring this curiosity through its effects on remembering. Both introspective and behavioural evidence reveal that, when an acceptable answer to a question has been encountered and rehearsed, curiosity is reduced to a subthreshold value. But the higher the drive before such reduction, the greater the *amount of reinforcement* or quantity of drive-reduction (K). And according to Hull's (1943, 1952) postulates, the probability of a response occurring on future occasions increases with reaction-potential ($^{S}E^{R}$), which in its turn increases with K. It follows that those questions which originally aroused most curiosity are most likely to be answered correctly if they are encountered again after the answer has been presented to the subject, and we can use the probability of recall as a measure of curiosity. An additional measure depends on the fact that subjects are likely to have learned to respond with tacts to their own internal stimuli (Skinner, 1953, 1947), although less accurately than to external stimuli. They can accordingly be instructed to indicate which questions arouse the greatest desire to know the answer.

We have therefore arrived at the hypothesis that curiosity is aroused in a subject when a question is put to him, whether by himself or by an external agent. Some component (s_{mD}) of the response-produced stimulation resulting from the meaning of the question (r_m) is assumed to act as a drive-stimulus. And we can see that the intensity of this drive-stimulus, which will in its turn depend on the amplitude of the response (r_{mH}) that produces it, will be one of the most important variables affecting the drive strength of the curiosity.

There is some experimental evidence for the curiosity-inducing role of questions, but it is also borne out by everyday experience. Many celebrated thinkers have been stimulated to a lifetime's meditation simply because they thought of new questions about matters that ordinary men have taken for granted. Similarly, the skilful lecturer excites curiosity in his audience by putting questions to them, perhaps about familiar phenomena, which it has never occurred to them to ask themselves.

However, the factors mentioned so far do not adequately explain the most striking cases of curiosity-arousal, those concerning the strange, the unusual, the puzzling. To attempt an explanation of this side of human nature, we shall have recourse to another variable, *conflict*.

III. THE ROLE OF CONFLICT

After the necessary preliminary phase of considering over-simplified situations, in which either only one response-tendency or motive is active or else one response-tendency or motive is so much stronger than others as to be virtually alone in its influence, psychological theory had to turn to more realistic situations where there are factors in competition. Even an elementary treatment of trial-and-error learning (Hull, 1930) forces us to consider the process whereby one response overcomes alternative ways of reacting, but special phenomena result when competing tendencies are fairly evenly matched in strength. The study of such phenomena was begun by Lewin (1935) and then carried further on both theoretical and experimental planes by Miller (1944, 1951) and his associates. Dollard & Miller (1950) have shown how the behaviour-theory of conflict can be extended to embrace the main effects ascribed to conflict by Freud, while Hull has endeavoured to reveal its roots in the basic principles of learning (1952, chap. 8). A theory of emotion, based on the assumption that conflict (F) is in itself drive-producing, is an important recent development for which Brown & Farber (1951) are responsible and there are various observations from experiments with rats that tend to confirm this assumption (Finger, 1941; Lowell, 1952; Miller & Stevenson, 1936).

A rather different recent emphasis on conflict has come from Hebb (1946, 1949). This is particularly deserving of mention here, because it involves the central processes intervening between stimulus and response, and that is precisely where we must seek the kernel of curiosity. Behaviour, in Hebb's view, depends on the intricate and nicely timed co-operation between 'cell-assemblies' in the cerebral cortex. If the timing goes awry, or if the processes ('phase-sequences') in the cortex otherwise interfere with one another, disruption will result. Some phase-sequences require the support of externally initiated sensory processes, and if these are not forthcoming, as when something familiar with an unexpected feature is perceived, disruption is once again a likely outcome. This disruption, which leads to a diffuse and disorganized release of energy, is what, according to Hebb, lies behind emotion. His principal illustration is his description of the fear induced in chimpanzees by surprising sights, but it is easy to see that these sights might instead have aroused curiosity in slightly different conditions. In his treatment of perceptual learning, Hebb describes how repeated exposure to a complex of stimuli builds up integrated and organized patterns of activity in the cortex, and thus conflict is eliminated as the unfamiliar becomes familiar. If we admit the possibility that the curiosity aroused by unusual perceptions has something to do with conflict, then the elimination of this conflict by exploration and the consequent drive-reduction might well play a part in perceptual curiosity. If we then extend these ideas to the autonomous processes which are the result of prior learning but can later run off in the absence of the corresponding environmental events, we can readily imagine how strange and puzzling thoughts or concepts may likewise involve conflict, and the acquisition of knowledge may mean the formation of new structures which obviate this interference. If conflict is a drive, the reduction of conflict will be reinforcing, and it will provide the explanation for the reward-value of investigating things that are puzzling and the learning of knowledge resulting from this investigation. Epistemic curiosity also will thus be attributable in many cases to a similar mechanism.

Hebb's concepts are physiological and refer to neural processes. But since these processes are at present not observable and serve merely as devices for explaining what can be observed, they are best regarded as intervening variables. It should not be difficult therefore to translate them into behavioural terms. The preference for purely behavioural terms may be justified as more than a matter of verbal taste, since these terms point to relationships between the sort of conflict under discussion and other areas in behaviour theory, including other forms of conflict.

We can begin our inquiry into the conflicts affecting trains of thought by recalling Miller's (1944) list of ways in which responses may be incompatible and therefore conflict. Sometimes the incompatibility is physical and innate, like that between approaching and avoiding the same object. But at other times, the conflict is learned. The responses are not inherently antagonistic, but learning has made the organism unlikely to perform both simultaneously or in close succession. This means that the response-produced stimulus (s_1) resulting from the first response (R_1) has become conditioned to a response (R_{-2}) which is physically incompatible with the second (R_2). If R_{-2} is stronger than R_2, the latter will thus be inhibited. When we extend these notions to symbolic responses, it is clear that physical incompatibility will not be of major importance. It may be that certain perceptual responses are innately incompatible, so that the conflict between an expectancy and a perception, for example, may be unlearned. But most of the antagonisms between symbols are almost certainly the results of learning, which trains us not to apply two particular words to the same object or combine two particular concepts (r_m) in the same complex. Thus a thought or a perception may conflict with past experience by incorporating two elements previously learned as incompatible. There are two cases of learned incompatibility that are particularly likely to affect trains of thought. One is the learned incompatibility between *contradictory* beliefs, which enables us to recognize and avoid fallacious symbol-sequences. The other is the conflict between the learned fear of *irrelevance* and the tendency to perform an irrelevant verbal response. As we have already noted, the Würzburg school were impressed with the way in which the determining tendency kept thought in the right direction by excluding irrelevant associations. Modern learning theory (Dollard & Miller, 1950) leads us to the conclusion that this happens because the emergence of an irrelevant thought evokes a learned drive and its inhibition reduces the drive and is consequently reinforced. The existence of a strong 'drive to be relevant', at least in our culture, can be demonstrated in almost any session of psychotherapy with a neurotic. Long before his free association has led him to touch on anything delicate, he shows resistance, due to his previous training to speak relevantly and coherently. It takes several sessions before he can flit inconsequentially from one topic to another or expatiate on matters that seem unconnected with his symptoms, as he must if he is to obey the 'basic rule'.

In order to show how conflict can affect curiosity, it is helpful once again to take a concrete example of a question: 'what crops do some ants cultivate in underground "farms"?' We have already introduced the assumption that a question arouses concepts (r_m), including some which are drive-producing. But, in addition, we can indicate, with the aid of our illustrative question, four stages, at any or all of which conflict may occur to bring about an increment of curiosity-drive:

(1) The question itself may evoke concepts which past experience and instruction have made incompatible for the subject. In the case of our example, learned conflict may well exist in a zoologically naïve person between all concepts relating to farming and all relating to subhuman animals. This is what happens when the designator of the question is said to be 'surprising' or 'unexpected' or 'strange' or 'puzzling'. Instead of such a question being put to the subject by an outside source, he might have come across a stimulus-situation arousing perceptual responses which he has been trained to regard as incompatible. This might lead him to take a closer look (perceptual curiosity) or to formulate such a question himself.

(2) Even if the question itself did not imply any surprising fact (e.g., 'how does the starfish eat?') conflict may well arise immediately after its formulation, if the answer is unknown. If the answer is known, then there will be a response already learned to the stimulus-pattern produced by the question, and this response will probably occur without delay. But if the answer is unknown, then we must turn to behaviour theory for some hints as to what might ensue. If the organism is confronted with a combination of stimuli to which it has learned no response, we shall not expect it to do just nothing, unless it is a very primitive or young organism indeed. As Hebb (1949) has aptly reminded us, even the most novel situation for an adult rat or human being is built up of elements which resemble some things he has met many times before. Therefore, if a new question is the stimulus-situation, the responses that we can expect to occur are those which are aroused by stimulus generalization from similar patterns or elements. The strongest will be those conditioned to patterns consisting of some identical and some slightly different elements, as compared with the present pattern, or else those conditioned to single elements. For

example, if no responses have been learned to a combination of 'ants' and 'farming', we can expect responses to occur which are associated with 'ants' plus some other activity, or 'farming' plus some other animal, or with 'ants' or 'farming' alone. This follows from Hull's treatment of stimulus generalization and of patterning (1952, chaps. XII, XIX). However, the trains of thought leading out from 'ants' and from 'farming', etc., are likely to be of comparable strength and incompatible. So, here again, conflict may add to the drive-strength.

(3) As these associative processes continue, they are likely to lead, in the absence of other trains of thought, to some which are irrelevant to the motivation to answer the question. This will produce the type of learned conflict to which we have already drawn attention.

(4) Finally, in all probability, concept-patterns will be reached, which are recognized as possible answers. If one is strong and the others are not, then that one will be accepted and learned, and there the process will end. But it may very well be that the subject is faced, either through his own cogitations, or through the intervention of some external agency (as in 'multiple-choice questions'), with a number of possible answers which seem about equally plausible. In that case, conflict between them is to be anticipated. Moreover, if any answer is of such a strength that tendencies to accept or reject it are more or less equal, we shall have another source of conflict, reminiscent of the approach-avoidance conflict. The conflict characteristic of this fourth stage is, of course, particularly prominent in the 'yes-or-no question' (or the 'which—question'). The drive to have the answer will be strongest, according to our expectations, if the tendencies to say 'yes' and 'no' are about equal.

Now, the drive produced in these various ways by conflict can only rightly be called 'curiosity' or a 'drive to know' if it is reduced by the process of knowledge-rehearsal. We must then see how exposure to concepts which are acceptable as the correct answer might lead to conflict-reduction. There are in fact three ways in which this might happen, corresponding to our four cases of conflict as follows:

(1) The answer, by implying that two concepts formerly regarded as incompatible need not be so, may inhibit learned conflict. This would mean that the S–R bond between s_1 and R_{-2}, to use Miller's (1944) notation, would be inhibited by the action of verbal stimuli (Dollard & Miller, 1950). To revert to our example, the subject would cease to find surprising the idea that ants engage in a kind of farming.

(2) and (3) The answer may reduce the conflict derived from irrelevant, generalized trains of thought by evoking a new response-sequence which is strong enough to crowd them out and prevent them from arising. Thus, our subject, having been told something of the activities of harvesting ants, will in future, when confronted with a pattern of concepts combining 'ants' with 'farming', be led off along trains of thought peculiar to harvesting ants and capable of overcoming irrelevant digressions.

(4) The answer may reduce conflict by strengthening one competing response and weakening others, thus reducing the equality between them. This happens when we are made to believe that one of our suspected answers is right and the other is wrong. There is henceforth, therefore, no appreciable competition between them.

Since conflict (F) is an intervening variable, we must ask what variables affect its magnitude. Brown & Farber (1951) postulate that F increases with (1) the absolute strength of the competing tendencies and (2) the equality between them. There is some experimental evidence for the first of these hypotheses. Sears & Hovland (1941) found that avoidance-avoidance conflicts provoke more blocking when the competing responses are stronger. The assumption that the symptoms of conflict only occur when the conflicting tendencies are about equal is confirmed time after time in Miller's (1951) experiments. Nevertheless, we are obliged to hypothesize that the intensity of the conflict-drive depends also on two more variables, which have been little studied, but which are of prime importance for conflict between thought-processes. These are (3) the number of conflicting response-tendencies and (4) the degree of incompatibility between them. A few remarks on each of these are in order.

First, Miller's treatment of conflict has been confined to cases where only two tendencies are in competition. For that matter, Lewin's, Freud's and Hebb's theories do not consider fully the possibility of there being more than two. Pioneering studies have of necessity to begin with the simplest cases, and there may be special reasons why dual rather than higher-order conflicts are the most frequent. When the theory of higher-order conflicts comes to be attacked it is likely to be found that two or more tendencies among those competing will usually have more in common than others, so that they form an alliance, and the situation resolves itself into what is virtually a dual conflict. It is interesting to note how this happens in Freudian theory. There are three components of the personality—id, ego, superego—but neurotic conflicts take the form *ego + superego* vs. *id* or *ego* vs. *id + superego* (Fenichel, 1945, p. 132). However, in the higher mental processes, higher-order conflicts may be common. Several trains of thought may be leading symbol-sequences in many incompatible directions at once, just as, in Hull's treatment of learning (e.g. 1930), one stimulus may arouse three or more reaction-potentials at once. It seems reasonable to assume that if the number of conflicting tendencies is increased, all things being equal, the severity of the conflict will increase.

Secondly, both Hull (1952, corollary xiv) and Miller (1944) speak of cases where the incompatibility between response-tendencies is absolute. But, as we have elsewhere mentioned (Berlyne, 1951, p. 144), even with innate reflexes intermediate degrees of incompatibility and partial interference are known. Learned conflict is likely to make even more indispensable the concept of degree of incompatibility, since the $s_1 \rightarrow R_{-2}$ bond may have various strengths. Subjects may, for instance, be surprised to different extents to hear of ants farming.

Our theory may link up with certain threads in contemporary social and developmental psychology. Many writers have been showing (Cantril, 1941, etc.) that one of the most distressing plights for human beings is not to know or to understand a state of affairs, particularly if it is important for their security or contrary to their expectations. One of society's most vital functions is to provide norms and frames of reference for the evaluation of new contingencies. If these are lacking or inapplicable to an unprecedented crisis-situation, people will be prone to accept and spread rumours or to succumb to suggestion. Many of the rumours and fanciful stories accepted by them in this gullible mood are the very reverse of reassuring. Indeed, they often show a tendency to believe in the most alarming prospects they can imagine, so that this desire to have an explanation available does not seem to be reinforced by fear-reduction. But the principal drive behind it may well be the conflict-drive, produced by uncertainty, which, as many wartime phenomena showed, is often more agonizing than realistic anticipation of unpleasantness. When there is a perplexing situation, raising an unanswered question, we predicted from the principles of behaviour theory that responses due to generalization from similar stimuli would occur. This casts light on the inveterate proclivity, so often observed in children and in primitive peoples, to interpret in terms of familiar phenomena those things that are of moment to them but they have little objective information about, which gives their speculations an 'artificialistic' and 'animistic' stamp (Piaget, 1931). Similarly, when irregular figures are exposed in a tachistoscope, Bartlett (1932) reports that subjects evince an 'effort after meaning' which leads them, as their immediate reaction, to relate the figures to something familiar.

Our theory of curiosity implies that patterns will be most curiosity-arousing at an intermediate stage of familiarity. If they are too unlike anything with which the subject is acquainted, the symbolic response-tendencies aroused will be too few and too feeble to provide much conflict, while too much familiarity will have removed conflict by making the particular combination an expected one. Once again, we have a prediction that accords with the observations of psychologists working in several fields. McDougall (1908) contended that something with an intermediate degree of familiarity is the adequate stimulus for curiosity. Hebb (1949) expresses the belief that the mental processes corresponding to cortical processes ('phase-sequences') will be most rewarding, and therefore most likely to occupy the subject, when the phase-sequences are in the course of being built up. The cell-assemblies and phase-cycles will then be in existence, but they will not yet have been moulded into a firm unity. Freudians (Fenichel, 1945, p. 45) attribute many of the play-activities of children to the discovery that they are now able 'to overcome without fear a situation that formerly would have overwhelmed (them) with anxiety'. But this 'functional pleasure' will be attached to a given activity only during the comparatively short interval between coming to fear it and triumphing over the fear. Piaget (1931, 1945) writes of the apparently useless actions that infants are prone to indulge in repeatedly (primary, secondary and tertiary circular reactions); but each of these actions dominates a child's activity for a brief period while he has discovered the ability to produce the effects in question but has not yet assimilated them to the point where they cease to be interesting.

So far our discussion has concerned itself with postulated and unobservable central processes, which are represented by intervening variables. In order to reveal the empirical content of our theory and the testable predictions that can be derived from it, it will be necessary to state the relations between our intervening variables and their observable antecedents and consequents. This will be done in a forthcoming article in this *Journal*, which will also report some experimental data tending to confirm our hypotheses.

IV. SUMMARY

Human 'epistemic curiosity', to be distinguished from the 'perceptual curiosity' that is found in lower animals as well as in human beings, is defined as a drive reducible by knowledge-rehearsal. An analysis of epistemic curiosity, using a behaviour-theory approach, is presented, its principal features being (1) an account of questions as 'thematic probes' which evoke drive-producing meaning-responses, and (2) the attribution to learned conflict of the curiosity aroused by strange, surprising or puzzling situations or questions.

This article is part of a dissertation submitted to the Faculty of the Graduate School of Yale University in partial fulfilment of the requirements for the degree of Doctor of Philosophy. The author is indebted to Dr C. I. Hovland, thesis adviser, and to Drs I. L. Child and N. E. Miller, members of the thesis committee, for their valuable help. The project was carried out as part of the Yale Communication Research Programme, which is directed by Dr Hovland and financed by a grant from the Rockefeller Foundation.

References

Abraham, K. (1927). Restrictions and transformations of scoptophilia in psychoneurosis: with remarks on analogous phenomena in folk psychology. (In *Selected Papers*. London: Institute of Psychoanalysis, and Hogarth.)

Ach, N. (1910). *Über den Willensakt und das Temperament*. Leipzig: Quelle and Meyer.

Bartlett, F. C. (1932). *Remembering*. Cambridge University Press.

Berlyne, D. E. (1950). Novelty and curiosity as determinants of exploratory behaviour. *Brit. J. Psychol.* li, 68–80.

Berlyne, D. E. (1951). Attention, perception and behaviour theory. *Psychol. Rev.* lviii, 137–46.

Berlyne, D. E. (1954). Knowledge and stimulus response psychology. *Psychol. Rev.* (in the Press).

Bloch, B. & Trager, C. L. (1942). *Outline of Linguistic Analysis*, pp. 52. Linguistic Society of America.

Brown, J. S. & Farber, I. E. (1951). Emotions conceptualized as intervening variables—with suggestions toward a theory of frustration. *Psychol. Bull.* xlviii, 465–95.

Cantril, H. (1941). *The Psychology of Social Movements*. New York: Wiley.

Dollard, J. & Miller, N. E. (1950). *Personality and Psychotherapy*, pp. 119–20. New York: McGraw Hill.

Fenichel, O. (1945). *The Psychoanalytic Theory of Neurosis*. New York: Norton.

Finger, F. W. (1941). Quantitative studies of 'conflict': I. Variations in latency and strength of the rat's response in a discrimination-jumping situation. *J. Comp. Psychol.* xxxi, 97–127.

Freud, S. (1905). *Three contributions to the Theory of Sex*. (In Brill, A. A. (ed.) *The Basic Writings of Sigmund Freud*. New York: Modern Library, 1938.)

Hebb, D. O. (1946). On the nature of fear. *Psychol. Rev.* liii, 259–76.

Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley.

Hull, C. L. (1930). Simple trial-and-error learning; a study in psychological theory. *Psychol. Rev.* xxxvii, 241–56.

Hull, C. L. (1943). *Principles of Behavior*. New York: Appleton-Century.

Hull, C. L. (1952). *A Behavior System*. New Haven: Yale University Press.

Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt Brace.

Lewin, K. (1917). Die psychische Tätigkeit bei der Hemmung von Willensvorgängen und das Grundgesetz der Assoziation. *Z. Psychol.* lxxvii, 212–47.

Lewin, K. (1922). Das Problem der Willensmessung und das Grundgesetz der Assoziation. *Psychol. Forsch.* i, 65–140, 191–302.

Lewin, K. (1935). *A Dynamic Theory of Personality*. New York: McGraw Hill.

Lowell, E. L. (1952). Experimental studies of the motivational effect of conflict. *Amer. Psychologist*, vii, 253–4 (Abstract).

McDougall, W. (1908). *An Introduction to Social Psychology*. London: Methuen.

Miller, N. E. (1944). Experimental Studies of Conflict. (In Hunt, J. McV. (ed.) *Personality and the Behavior Disorders*. New York: Ronald.)

Miller, N. E. (1951). Comments on theoretical models—illustrated by the development of a theory of conflict behavior. *J. Personal.* xx, 82–100.

Miller, N. E. & Stevenson, S. S. (1936). Agitated behavior of rats during experimental extinction and a curve of spontaneous recovery. *J. Comp. Psychol.* xxi, 205–31.

Montague, W. P. (1936). *The Ways of Knowing*. London: Allen and Unwin.

Morris, C. W. (1946). *Signs, Language and Behavior*. New York: Prentice-Hall.

Mowrer, O. H. (1939). A stimulus-response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.* xlvi, 553–65.

Mowrer, O. H. (1950). *Learning Theory and Personality Dynamics*. New York: Ronald.

Mowrer, O. H. & Ullman, A. D. (1945). Time as a determinant in integrative learning. *Psychol. Rev.* lii, 61–90.

MOWRER, O. H. & VIEK, P. (1948). An experimental analogue of fear from a sense of helplessness. *J. Abn. (Soc.) Psychol.* XLIII, 193–200.

OSGOOD, C. E. (1952). The nature and measurement of meaning. *Psychol. Bull.* XLIX, 197–237.

PIAGET, J. (1931). Children's philosophies. (In Murchison, C., *Handbook of Child Psychology*. Worcester, Mass.: Clark University Press.)

PIAGET, J. (1945). *La formation du symbole chez l'enfant*. Neuchâtel: Delachaux and Niestlé. (*Play, Dreams and Imitation in Childhood*. New York: Norton, 1951).

PIAGET, J. (1947). *La psychologie de l'intelligence*. Paris: Colin. (*The Psychology of Intelligence*. London: Routledge, 1950.)

SEARS, R. R. & HOVLAND, C. I. (1941). Experiments on motor conflicts. I. Determination of mode of resolution by comparing strengths of conflicting responses. *J. Exp. Psychol.* XXVIII, 280–6.

SKINNER, B. F. (1953). *Science and Human Behavior*. New York: Macmillan.

SKINNER, B. F. (1947). *Verbal Behavior*. William James Lectures, Harvard University, 1947. (To be published.)

ULLMAN, A. D. (1951). The experimental production and analysis of a 'compulsive eating symptom' in rats. *J. Comp. Physiol. Psychol.* XLIV, 575–81.

WERTHEIMER, M. (1945). *Productive Thinking*. New York: Harper.

(*Manuscript received* 14 *December* 1953)

A NOTE ON ASHANTI NAMES AND THEIR RELATIONSHIP TO PERSONALITY

By G. JAHODA

University College of the Gold Coast

I. *Introduction* (p. 192). II. *Background of traditional beliefs* (pp. 193). III. *Methods used* (pp. 193–194). IV. *Results and conclusion* (pp. 194–195). *References* (p. 195).

I. Introduction

The present paper originated with a discussion at Kumasi, Ashanti, in which mainly teachers and welfare officers took part. The question at issue was the development of character in children. Several of the people present maintained that the day of the week on which a person is born determines what type of character will emerge. The writer, who had not long been in Africa at that time, was politely sceptical. But one headmaster, whose previous contributions had proved him to be an intelligent and well-informed man, claimed that he had put this matter to the test. It appeared that he had consulted his records of corporal punishment and had found that those born on Wednesdays were by far the most frequently in trouble; this accords with popular beliefs.

The writer was sufficiently impressed by this statement to consider the possibility, not of course that the day of birth determines character, but that prevailing beliefs in this matter may do so, especially when they are held by educated as well as by illiterate Ashanti. Over half the people present at the meeting were firmly convinced that differences of the kind indicated are a fact; the remainder, whilst knowing about these beliefs, were doubtful; only two out of sixteen denounced them as superstitions.

The view that the expectations of other people influence the development of a person's character has been held by many psychologists. For instance McDougall (1936) wrote:

> (the child) gets his idea of his self in large part by accepting the ideas of himself that he finds expressed by those about him. The process is well illustrated by the case of the unfortunate child who is constantly scolded and told that he is a naughty boy. Under these conditions the normal child very soon accepts these oft-repeated suggestions, learns to regard himself as a naughty boy, and plays the part thus assigned to him. Similarly, if he finds himself constantly regarded as clever, or irresistibly charming, or in any other light, he can hardly fail to regard himself in the same way, and the idea of his self moulded in this way by his social environment affects his conduct accordingly.

Others have followed McDougall, for example, Murphy, Murphy & Newcomb (1937), but there appears to be no supporting empirical evidence. The reason is perhaps to be sought in the fact that parents' and other people's attitudes towards children vary greatly and irregularly in Western Society, thereby making the task of tracing specific effects a hopelessly complicated one. In the present case an element of uniformity is introduced by the existence of socially determined beliefs about certain defined categories of children. It therefore becomes possible to use a quantitative approach, provided satisfactory criteria of conduct or personality can be found.

II. BACKGROUND OF TRADITIONAL BELIEFS

Before describing the methods used in the investigation, a brief outline of beliefs about names is necessary. Every child in Ashanti and other *Akan* areas receives a name from the day on which he is born. The name thus derived is called *kradin* or soul name; it refers to the day on which the soul is 'washed' or purified (cf. Rattray, 1927). These names are listed in Table 1.

Table 1. *Ashanti 'day names'*

Day of the week	Boys	Girls
Sunday	*Kwasi*	*Akosua*
Monday	*Kwadwo*	*Adwoa*
Tuesday	*Kwabena*	*Abenaa*
Wednesday	*Kwaku*	*Akua*
Thursday	*Yaw*	*Yaa*
Friday	*Kofi*	*Afua*
Saturday	*Kwame*	*Amma*

There is no evidence that any significance is attached to girls' names. With boys the beliefs about character traits corresponding to each 'day name' show some regional variation, but there is widespread consensus about two of them: a *Kwadwo* is supposed to be quiet, retiring and peaceful; by contrast a *Kwaku* is held to be quick-tempered, aggressive and a trouble maker. Similar conceptions are implicit in the appellations relating to these days. That for Monday is *Awo*: 'It is the day of the moon or the vesper which is "cool"; the day of peace' (Danquah, 1945). Wednesday's address is *Awuku*; the root *wu* in this means 'to die'; *di awu* is 'to commit a murder' (Christaller, 1933). Furthermore, a phrase is current about *Kwaku babone*, the bad Wednesday boy; one informant told me that during the War people in his village used to talk about *Kwaku* Hitler.

It should perhaps be mentioned that day names merely indicate that a particular type of *kra* has entered the body; such a *kra* produces a disposition towards certain kinds of behaviour, but the disposition may to some extent be counteracted and does not amount to a fixed destiny.

The naming customs described are based upon a magico-religious system of ancestor-worship which still affects Ashanti life in numerous ways, in spite of widespread, although often superficial, conversion to Christianity. Literates, who are usually Christians (education having been until recent years largely in the hands of missions) represent a minority of about 6% of the population. Practically all of them continue to give their children day names and very few are entirely divorced from the older ideas.

III. METHODS USED

It is clear, therefore, that these beliefs are powerfully entrenched in tradition, and likely to produce a set of expectations concerning the ways in which a *Kwadwo* or a *Kwaku* will behave. A direct check of this hypothesis was unfortunately impracticable for someone who is not a fulltime field worker; it would have involved prolonged and intensive study of differential attitudes, both overt and covert, towards those born on the various days of the week. More superficial inquiry did reveal that phrases like 'A *Kwaku* naturally WOULD do this kind of thing' are quite commonly used.

Turning from process to product, namely the resulting personality differences, other differences are encountered. It is unlikely that such differences could be readily detected, and suitable criteria established, with available personality 'tests' or techniques; in any case these are as yet unproven in Africa.

A rather crude test, similar to the one used by the headmaster, is however possible, which has the advantage of providing fairly clear-cut behavioural criteria. It consists in examining the Juvenile Court records for the region to ascertain the frequency of conviction and type of delinquency committed by boys born on different days of the week. This task was accordingly undertaken, without any sanguine hopes of achieving significant results.

As a preliminary step the names of all boys in ten Kumasi schools were obtained. This was done to ensure that no general bias was present in the distribution of *kradin*; for it is conceivable, for instance, that in the case of births occurring around dawn the 'good' rather than the 'bad' day might be chosen; it has to be understood of course that there may be a variety of reasons why a particular day should be thus regarded.

Then the delinquency records were consulted for the period between September 1948 (when the Juvenile Court started working) and June 1953. Boys belonging to tribes other than Ashanti were excluded. Offences were divided into those against the person (e.g. assault, fighting, causing wound), against property (e.g. stealing, receiving, fraud) and 'other offences' (e.g. traffic offences, unlawful assembly, wandering). Recidivists were classified according to their first offence. In spite of repeated careful checking aliases may have been a source of error; but there is no reason to suppose that their incidence would affect some day names more than others; furthermore, there were indications that they occurred only rarely, so that their total effect was almost certainly negligible.

IV. RESULTS AND CONCLUSION

The heading 'Other' in Table 2 refers to those cases where no information about *kradin* was available; this happened more often with the delinquents, when existing data had to be used. School-teachers were specially requested not to omit them.

Table 2. *Distribution of names: ordinary schoolboys and delinquents*

Ashanti names...	*Kwasi*	*Kwadwo*	*Kwabena*	*Kwaku*	*Yaw*	*Kofi*	*Kwame*	Other*	Total
Corresponding days ...	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday		
Schoolboys	179	157	151	164	163	179	188	73	1254
Delinquents	47	23	42	45	49	63	63	114	446

Test of significance D.F. $=6$, $\chi^2=15{\cdot}75$, $P=0{\cdot}02$.

* The category 'Other' was omitted in computing χ^2.

It will be seen that there is a significant difference between the two distributions, which is mainly accounted for by the small proportion of *Kwadwos* among the delinquents: this alone contributes a proportionate square of 10·13 out of a total χ^2 of 15·75. *Kwadwos* thus appear to live up to their reputation of being quiet and well-behaved as compared with boys born on other days.

On the other hand, the supposedly difficult *Kwakus* (numbering 45) come close to the expected figure (47·43) as far as all types of delinquency are concerned. However, if one considers only offences against the person, which are presumably the most direct expression of an aggressive disposition, the picture fits again. For *kradin* other than *Kwaku* the absolute number of such offences ranges from 2 (*Kwadwo*) to 8 (*Yaw*), with a mean of 4·8; for *Kwaku* the number is 12 and this, as is apparent from Table 3, represents a significantly higher proportion than that found in all the rest taken together.

Table 3. *Offences against the person*

Type of name	Offences against the person	All other offences	All offences
All except *Kwaku* (including 'other')	43	358	401
Kwaku only	12	33	45
All delinquents	55	391	446

Test of significance D.F. = 1, $\chi^2 = 9{\cdot}52$, $P = 0{\cdot}002$.

The results here presented are consistent with the hypothesis that Ashanti beliefs about a connexion between personality characteristics and day of birth may be effective in selectively enhancing certain traits which otherwise may have remained latent. The evidence is of course indirect and as such not conclusive. Nevertheless, the correspondence appears too striking to be easily dismissed. It is probable that for many Ashanti there is some truth in the saying *nomen est omen*.

The writer wishes to acknowledge with thanks the assistance of the Regional Director of Education, Ashanti, and the Probation Officer, Kumasi. He is greatly indebted to his colleague Mr Joseph Kwabena Nketia, Research Fellow in African Studies, without whose contribution this study could not have been undertaken.

REFERENCES

CHRISTALLER, J. G. (1933). *Dictionary of the Asante and Fante Language.* Basel: Evangelical Missionary Society.
DANQUAH, J. B. (1945). *The Gold Coast Akan.* London: Lutterworth Press.
McDOUGALL, W. (1936). *An Introduction to Social Psychology*, 25th ed., p. 160. London: Methuen and Co.
MURPHY, G., MURPHY, L. B. & NEWCOMB, T. M. (1937). *Experimental Social Psychology*, p. 211. New York: Harper and brothers.
RATTRAY, R. S. (1927). *Religion and Art in Ashanti.* Oxford: Clarendon Press.

(*Manuscript received* 23 *February* 1954)

EVIDENCE OF IMPROVEMENT IN POETIC JUDGEMENT

BY J. N. BRITTON

Institute of Education, University of London

I. *Introduction* (p. 196). II. *Aim* (p. 196). III. *Procedure* (pp. 196–198): (i) *Preparing the test* (pp. 196–197); (ii) *Setting the test* (p. 197); (iii) *Methods of analysis* (pp. 197–198). IV. *Results* (pp. 198–205): (i) *Performance of the eighteen experts* (pp. 198–199); (ii) *Critical comment upon the experts' performance* (p. 199); (iii) *Performance of the whole sample of* 221 *raters* (pp. 199–200); (iv) *Analysis by groups of raters* (p. 200); (v) *Critical consideration of four points arising* (pp. 200–202); (vi) *The test repeated* (pp. 202–203); (vii) *A comparison of average orders* (pp. 204–205); (viii) *The convergence of preference* (p. 205). V. *Summary and conclusions* (pp. 206–207). *References* (p. 207). *Appendix: details of test items and examples of counterfeit poems* (pp. 207–208).

I. INTRODUCTION

Experience seems to suggest that changes of opinion concerning works of art include among them a number that are the effect of maturing judgement or 'improvement' in taste. That is to say that a number of these changes are such that they would be likely to move in one direction (say to promote item *A* above item *B*), unlikely to move in the opposite direction (to promote *B* above *A*), and extremely unlikely to be reversed, once made in what, by hypothesis, is the right direction. If this is the case it should be possible, among a sufficient number of observations, to discern a tendency for opinions to converge upon certain items at the expense of certain others. This hypothesis is tested here in the field of poetry.

In a previous experiment the writer (Britton, 1953), following Eysenck (1940), found clear evidence of the importance in poetic preference of the distinction between simple poetry and complex. 'Determiners' were prepared for this factor, as also for Eysenck's second bipolar factor—that distinguishing between poetry of restrained feeling and poetry of 'great emotionality'—by pooling ratings carried out by experts for the criteria concerned. The material used consisted of thirty-one short poems by modern writers, selected as being of approximately equal esteem in the opinions of experts.

II. AIM

The purposes of the present experiment were: (i) to employ a 'test' to measure the response of a number of persons to 'goodness' and 'badness' in poetry; (ii) to repeat the test after an interval in order to try out the hypothesis that there would be a convergence of preference upon the good poems (i.e. an 'improvement' in poetic judgement); and (iii) to compare the influence of the distinction between good and bad poetry with the influence of the two bipolar factors referred to above.

III. PROCEDURE

(i) *Preparing the test*

There is an obvious danger here that doubt may be cast upon the validity of the distinction between good and bad poetry as reflected in the examples chosen. Can the badness and goodness of a test poem be established as more than a matter of the experimenter's opinion? The pooled opinions of experts provide satisfactory measures of fairly complex criteria, but so simple a means of setting up objective aesthetic value itself would be unacceptable.

A solution was found in this case by avoiding any kind of hierarchy in the material and using instead a discrete, dichotomous arrangement. No attempt was made to distinguish degrees of goodness among the good poems: eight poems were taken from thirty-one used in a previous experiment, and it was assumed that in each the poet had something to communicate: that he had no intention in writing other than such as he hoped a reader would discern. Nor was any attempt made to distinguish degrees of badness: the bad poems were counterfeit or spurious, concocted by the experimenter with the deliberate intention of deceiving readers. They were made to look like poetry, but, *since they have nothing to communicate*, they are not.

The true poems were selected so that two poems represented each of the four poles of Eysenck's factors. Seven false poems were prepared in rough conformity with the same four categories. The following diagram illustrates the composition of the test: examples of the false poems and details of the true poems are given in the Appendix.

	True poem nos.	False poem nos.
Simple	5 and 12	1 and 3
Complex	2 and 11	4 and 14
Abandoned	6 and 10	7 and 13
Restrained	9 and 15	8

(ii) *Setting the test*

Subjects were asked to arrange the poems in the order of their preference and, wherever possible, add comments giving reasons for their assessment. Two hundred and twenty-one subjects took the test on the first occasion, eighteen of them being experts (English honours graduates, teachers of poetry), the remainder non-experts (including ten postgraduate English students, ten English honours undergraduates, forty training college students, eighteen Services personnel, and 110 sixth-form pupils, of whom sixty-four were taking arts courses and forty-six science courses).

A hundred and twenty of these took the test a second time after an interval of 4–6 months. These included ten postgraduate students, four undergraduates, forty training college students and sixty-eight sixth-form pupils, forty-two in arts and twenty-six in science.

Raters on the second occasion were asked to give their *present* reactions, 'which might conceivably differ from your previous ones'.

(iii) *Methods of analysis*

Preference ratings were allotted scores roughly in accordance with normal distribution as follows:

Score	+4	+3	+2	+1	0	−1	−2	−3	−4
Frequency	1	1	2	2	3	2	2	1	1

The significance of poem totals or category totals, above and below the mean of zero, was measured by Fisher's adaptation of Student's test for use with small samples (Fisher, 1936, pp. 125–6). The significance of changes from first to second rating was also calculated in this way, using differences between members of each pair of scores as themselves scores. Analysis of variance was employed to discover whether the true/false distinction could be seen to influence raters, and how its effect compared with that of the bipolar factors.

An analysis of the variance between and within poems was first carried out, and this was followed by an analysis of the variance between and within the categories 'true' and 'false', and between and within the four bipolar categories—these analyses being carried out on the whole series of fifteen poems. Significance was measured by comparing the variance between categories, in turn, with the variance remaining after that attributable to both sets of categories had been deducted.

When, to quote an example, the ratings of the eighteen experts were analysed in this way it was found that the variance between true and false poem groups was significant (in favour of true) at the 0·1% level, and the variance attributable to the four bipolar categories was also significant, but at the 1% level.

More interesting light was shed upon the bipolar categories by analysing separately the group of poems classified as simple or complex, and the group classified as abandoned or restrained. Variance between true and false poems was also calculated in each case. The eighteen experts, for instance, showed a preference for the complex poems which was significant at the 5% level, some apparent preference for the true poems which did not reach significance level, and, in the other group of poems, a preference at the 0·1% level both for the restrained and for the true poems.

This analysis was felt, however, to be inadequate as a means of assessing the influence of the bipolar factors. In studying the effect of the true/false distinction the main concern was with the effect in one direction: the power of attraction of true poetry. In assessing such a power, the choice of true poems by some raters would be rightly offset by the failure of others to prefer them. But in the case of the bipolar factors, by their very bipolarity, the attraction of, for example, complexity for some readers and the attraction of its opposite for others should count alike as part of the influence of the bipolar factor. A group of raters which was equally divided in this respect, however strongly its members were attracted and repelled by poems in the category concerned, would yield a non-significant result. Clearly this masking effect made the method unsuitable.

The idea was considered of reversing the direction of some scores so that all differences appeared in the same direction, and an adequate representation of the *extent* of the variance thus given. It was decided, however, that a similar indication of its extent could be given together with indication of *direction* if two analyses were carried out: one for those raters who preferred one category in the pair and a second for those who preferred the other. This was in fact carried out wherever two groups of opinion existed.

Finally, similar analyses (in two groups of raters where necessary) were carried out upon the scores for true poems only. It will be recalled that true poems were allotted to bipolar categories on the evidence of the experts, whereas false poems were not.

These methods were employed both to analyse the total variance for all raters, and also to analyse the variance of selected groups of raters: significances were sought, in terms of which the preferences of these groups might be characterized and distinguished.

IV. Results

(i) *Performance of the eighteen experts*

The experts showed a highly significant preference for the true poems as a class (mean true-poem score per rater, +4·056; s.e. 0·962). Three true poems (nos. 6 and 10, abandoned, and no. 12, simple) had negative totals, and three false poems (nos. 4 and 14, complex, and no. 1, simple) had positive totals, though none of these reached the level of significance.

By analysis of variance it was found that the variance attributable to the true/false distinction, in favour of true, was very highly significant, and so also was that attributable to the four bipolar factors. Further analysis showed that the discrimination in favour of true poems operated principally among the abandoned/restrained poems, and was accompanied by a general and highly significant preference for the restrained poems. Among the simple/complex poems, on the other hand, there was no significant preference for the true poems, but both simple and complex poems had their adherents among the raters, preference reaching the highly significant level. This one-sidedness of the abandon/restraint distinction and two-sidedness of the simple/complex distinction tended to operate with all groups of raters.

(ii) *Critical comment upon the experts' performance*

A number of experts failed to achieve a positive true-poem total, and this raises an important issue. Is there any justification for expecting the experts to prefer the true poems (of whatever kind) to the false (of whatever kind)? The rating, after all, was one of liking and not of value.

There can, in the long run, be only one answer. A poem is not comprehended until some experience has been communicated to the reader by its words. Until it is comprehended it may be liked (or disliked) for many reasons, all of them secondary: the sound of the words may please, the ideas stated may be acceptable. If, finally, it becomes clear that the 'poem' has no experience to communicate, it can only be rejected as worthless or counterfeit.

A time factor, however, enters into the present situation. It takes time to discover a hoax, as it takes time to comprehend a poem: both remain 'potential' until they have made, or finally failed to make, some kind of impact upon the reader. It is suggested that wherever a poem did make an impact upon a reader it was either positively endorsed or positively rejected. Now the nature of the false poems is such that they would be less likely to make an impact than the true ones: 'making an impact' is a way of describing the process of synthesis as it begins to work in the reader's consciousness; and in such a process nothing succeeds like success. A negative result—no impact, no synthesis, no experience, no poem—would be likely to take longer to arrive at. It is, as it were, a process of elimination. It follows that the false poems (if they succeeded in deceiving at all) would tend to remain *sub judice* longer than the true poems.

(iii) *Performance of the whole sample of* 221 *raters*

Whereas poem totals in the experts' case showed no significant results in the wrong direction, the totals given by the whole sample show significant disapproval of true two poems (no. 2, complex, and no. 10, abandoned), and significant approval of four false poems (nos. 4 and 14, complex, no. 8, restrained, and no. 13, abandoned).

Analysis of variance of the whole sample demonstrated the importance of the bipolar categories as controlling preference (though the influence of the 'abandoned' category seemed to work in a negative direction usually): and the complete ineffectiveness of the true/false distinction. In the whole table summarizing results of the various analyses carried out, only one entry occurs in the true/false column: this indicates that 29% of the raters showed a significant preference for abandoned poems together with an only slightly less significant preference for the false poems in this group.

(iv) *Analysis by groups of raters*

Raters were grouped in three ways:

A. By maturity (i) Experts and students
  (ii) Second- and third-year sixth forms
  (iii) First-year sixth forms

B. By 'faculty' (i) Arts sixth forms
  (ii) Science sixth forms

C. By sex (i) Girls
  (ii) Boys

Of these seven subcategories:

(*a*) Only the experts and students showed in the analysis of variance any preference for the true poems as a class. The boys showed a significant preference, and the science students a highly significant preference, for the false poems.

(*b*) The division by maturity demonstrated very clearly the gradual change from preferring simple poems, in the earlier stages, to preferring complex poems in the later.

(*c*) The arts students distributed their favours fairly evenly among the bipolar categories (significant preferences, 56% pro simple, 44% pro complex: 38% pro abandoned, 62% pro restrained), whereas the science students showed the most one-sided distribution (74% pro simple, 26%—but not reaching significance—pro complex: 26% pro abandoned, 74% pro restrained).

(*d*) Girls' results differed very little indeed from those of boys.

(v) *Critical consideration of four points arising*

From a detailed study of these results, four matters are raised for special comment.

First, it was found that in all groups except the experts and students a preference for true simple poems existed side by side with a liking for false complex poems. (A large interaction term was found in carrying out a full analysis of variance of ratings for these poems.)

There is no need to explain at length the average rater's rejection of false simple poems: they were easily seen to be what they were—meaningless. Discrimination in favour of false complex poems is, however, a more difficult matter.

There is ample evidence, from group scores and from occasional comments, that all but the poorest raters tended to suspend judgement upon a poem they did not understand. This would have the effect of placing such a poem below those understood and liked and above those understood (or seen through) and disliked: in other words, on average, near the middle of the scale.

Other considerations, however, resulted in discrimination between the true and the false complex poems, the former being demoted from the neutral position. Study of comments shows that raters frequently gave lukewarm approval, or mild or trivial disapproval to the false complex poems (referring often to poetical language, familiar allusions, vague air of mystery, etc.): whereas many indignant attacks were launched against the tone or manner or apparent attitude of the writer in the true complex poems. It will be noted that in both cases (in default of knowledge of the poems' full meaning) it is incidental qualities or superficial appearances that arouse these reactions.

It is suggested that the originality itself, the bold intentions, of the true complex poems, brought about this resentment. For the poet to have his way with a reader in a poem of this kind would have meant a more *disturbing* experience for him than would be the case with a less complex or less original poem. When for any reason a reader is unable to respond to a strong stimulus of this sort a sense of frustration, irritation, resentment, would appear to be a more usual reaction than mere indifference.

The second matter is of a similar nature, and, once again, is probably the phenomenon underlying a fairly large interaction term found in analysing variance in some groups' ratings of the abandoned/restrained poems. There seems to have been among the less experienced raters a tendency to discriminate against true abandoned poems and yet distribute preferences fairly evenly among the remaining three categories (false abandoned, true restrained, false restrained). Why should a consistent preference for restrained poems be accompanied by a willingness to approve false abandoned poems?

Of the false abandoned poems, one (no. 7) dealt with love in its familiar form of love between the sexes: the other (no. 13) was a sentimental treatment of nature-and-friendship, a pathetic fallacy in the vein of Ella Wheeler Wilcox. The two true abandoned poems were both love poems in the accepted sense.

It is quite clearly a rejection of love poems and a strong preference for the nature poem that principally brought about the result under consideration. Two groups, however—the science forms and the boys—also prefer the false love poem to either of the true ones. One explanation will be offered for both tendencies.

It may be said of conventional symbols and stock images that, as 'objective correlatives', to use Eliot's phrase, they are useful and reliable.[1] No normal person can completely disguise from himself the fact that he is capable of comparatively strong feelings, and, however he may wish it to appear to others, he is at times willing or more than willing to respond to people, objects or experiences which excite such feelings and in some measure satisfy them. The stock situation in life and art, because it is already familiar, enables him to do so within strict bounds: it acts rather as a rein than as a spur. The unfamiliar stimulus of an original and powerful work of the imagination is always a threat to these self-imposed bounds and will often, it is suggested, be rejected for this very reason. This may account for the curious fact that habitual readers of a kind of writing that is regarded as 'sentimental' (or even 'sloppy') by experienced critics, themselves describe any original expression of strong feelings as sentimental or sloppy. The terms indicate equally a rejection, but the reasons in each case are, as has been suggested, very different.

If a tendency to cling to stock symbols and situations is a self-protective measure, it may well be particularly resorted to during phases of emotional instability; as, for example, during the period of adolescence.

It is part of a cultural education to teach a child to supersede the stock situation in what he reads, in order that art may so come to enrich experience. If education succeeds at all in this respect, the stock response (see Richards, 1935, pp. 15–16) would be more prevalent among those whose cultural training has been for any reason neglected. And this, unfortunately, one might expect to find more often in the science forms, especially among boys. Especially among boys, because for one thing it is traditional in this country to regard free expression of feeling as effeminate. The tradition may even be founded upon a real difference between the sexes, or may result from social divergences in the role of man and woman. Once again, if tradition and inclination combine to suggest caution to the boy, a part of his response is likely to be a prejudice in favour of the stock situation and the conventional symbol at the expense of genuine feeling new-made into poetry.

---

[1] 'The only way of expressing emotion in the form of art is by finding an "objective correlative"; in other words a set of objects, a situation, a chain of events which shall be the formula of that *particular* emotion, such that when the external facts which must terminate in sensory experience are given, the emotion is immediately evoked' (Eliot, 1934, p. 61).

The third point is that among the undergraduates as a group, and affecting also a number of training college students, there was found a strong vein of antipathy to simple poetry, whether true or false, and a corresponding preference for complex poetry, true or false. This effect will not surprise readers who have dealings with undergraduates, though the sharpness with which it appears in the figures and the intensity of some of the comments might. It is probably explained in part by the fact that one of the processes involved in becoming an experienced reader is that of acquiring the skill to comprehend more and more difficult writing; and that pride in this maturing process may be accompanied by impatience of any work which does not call forth the skill concerned. Or, more simply, it may be a matter of regarding as childish anything that a child might understand—a judgement appropriate perhaps to the transitional stage of adolescence.

Finally, comments gave abundant evidence of the operation of prejudice. It must be borne in mind that the process of judgement necessarily consists in referring data observed to standards already held, the fruits of past experience. It is when the fruits of past experience obtrude to hamper the observation of data that they may be called prejudices.

The commonest form of prejudice met was that of looking only for the recurrence of old and familiar pleasures: approving recollections of what is already thoroughly acceptable without asking for any new light to be shed upon it. 'I am an admirer of nature poetry and this poem reminds me of Keats's poem *To Autumn*', was one girl's comment upon poem no. 13.

There were, on the other hand, many extraordinarily precise preconceptions: as that 'never' in the last line of poem no. 12 should have been 'ne'er'; or that poems should have so many lines to a stanza and rhyme in such and such a way.

The strongest distortion seemed to come from a preconceived notion as to what a poet's *attitude* should have been towards his theme. Failure to evince the attitude expected resulted often in a charge of insincerity. There were widespread doubts as to Edward Thomas's sincerity in poem no. 12.

(vi) *The test repeated*

The test was set a second time in order to discover whether, among the changes observed, there was any tendency for preferences to converge upon the true poems. There was. Individual true-poem totals increased to give a difference of mean of 1·250 compared with S.E. of difference of means of 0·326: and a similar result was obtained by taking individual differences between first and second true-poem/totals for each rater and comparing the mean of this distribution with its S.E. ($\bar{x}=1{\cdot}250$; S.E. $=0{\cdot}331$). Changes of opinion in the group as a whole, therefore, concerning the poems of the test showed an improvement, a convergence upon the true poems at the expense of the false.

Table 1 represents graphically the changes made in assessment of each poem by the whole group of 120 raters. Significant variations from the mean are indicated, for each occasion, in the first and third columns: significance of difference of means, representing changes of opinion (increases being regarded as positive), form the middle column. It will be seen that the two true complex poems (nos. 2 and 11) and one true restrained poem (no. 9) show significant increases, while one true simple poem (no. 12) deteriorates significantly. Among the false poems there is only one significant change (at the lower level, under 5% probability)—the deterioration of poem no. 7 (abandoned).

Table 1. *Results of significance of means, and of difference of means, tests by poems* (*for* 120 *raters*)

(Single sign = under 5% probability; double sign = under 1% probability.)

True poems		First rating	Difference	Second rating
Simple	No. 5	(+ +)	.	(+ +)
	No: 12	.	− −	(− −)
Complex	No. 2	(− −)	+ +	.
	No. 11	.	+ +	(+ +)
Abandoned	No. 6	.	.	(−)
	No. 10	(− −)	.	(− −)
Restrained	No 9	(+ +)	+ +	(+ +)
	No. 15	(+ +)	.	(+ +)
False poems				
Simple	No. 1	(− −)	.	(− −)
	No. 3	(− −)	.	(− −)
Complex	No. 4	(+ +)	.	(+ +)
	No. 14	(+ +)	.	(+ +)
Abandoned	No. 7	(− −)	−	(− −)
	No. 13	(+ +)	.	.
Restrained	No. 8	(+ +)	.	.

It is suggested that the process of first suspending judgement on the complex poems and later, in the case of the true complex poems, coming down on the side of approval, would explain the entries in the table for poems 2, 11, 4 and 14.

The deterioration of poem no. 12 provides a useful warning: that it is a general tendency among the raters cannot be denied, and yet it is not in the direction looked for. The simplicity, the regular verse form and, to some degree, the sentiments expressed earned for this poem from a number of raters the comment 'Georgian'—a term that already has in it the opprobrium of one generation for the generation that immediately precedes it. The element of fashion, it is suggested, will often lead to sweeping changes of opinion in poetry: but, viewed broadly enough, these will be temporary and subservient to those ampler movements, slower but more certain, by which opinion would be found to converge upon genuine poetry.

It is gratifying to notice the comparative lack of change relating to false poems. Changes have to be caused, and the impact of the true poem would be more likely to incur change than the failure of the false poem (with no experience to communicate) to make an impact on the reader. Changes in position of the false poems are more likely therefore to be by gradual displacement.

(vii) *A comparison of average orders*

Candidates who took the test both times were divided into groups similar to those used to divide the whole sample of 221 raters, with the exception that on this second occasion there were two categories only in the division by maturity (second and third-year sixth forms: first-year sixth forms). Although significant changes in group poem-totals were fully examined, the more rewarding method of comparing performances proved to be by considering average orders of poems. Three standards were used as bases of comparison:

The average order of the whole sample of 221 raters.
The average order of the eighteen experts.
A skeleton 'order' in which true poems were given priority over false poems and all other differences ignored.

Comparisons with the first and second of these (converted into scores as described in §III (iii) above) were made by product-moment correlation: a similar index for comparison with the third was obtained by correlating the set of scores to be measured with a hypothetical set in which the numerical scores were a repetition of its own, but the signs were made positive for all true poems and negative for all false. The three measures intercorrelated thus:

	18 experts	*T/F* order
221 raters	0·43	0·06
18 experts	·	0·91

an indication of the variation in influence of the true/false distinction as between experts and the whole sample.

The average orders for 120 raters at first and second ratings reflect, by these standards, the improvement already shown to exist:

Correlation with ...	221 raters	18 experts	*T/F* order
First rating	0·91	0·36	0·23
Second rating	0·84	0·77	0·43

Breaking these results down into terms of groups of raters gives Table 2 below.

Table 2. *Characteristics of group average orders at first and second ratings*

Correlation with ...	221 raters		18 experts		*T/F* order	
Rating ...	1st	2nd	1st	2nd	1st	2nd
Second and third year	0·61	0·57	0·54	0·76	0·46	0·71
First year	0·87	0·93	0·06	0·63	−0·40	0·34
Arts forms	0·70	0·74	0·09	0·50	0·03	0·46
Science forms	0·66	0·80	−0·27	0·13	−0·57	−0·26
Boys	0·61	0·60	−0·13	0·23	0·06	0·06
Girls	0·77	0·83	0·10	0·60	0·00	0·26

Slight changes in the first (double) column indicate the fairly stable degree of general agreement or conformity among the raters, on either occasion or by any classification. Within this, however, lies considerable room for movement towards the more informed opinions of the experts (second column) and the objective true/false standard (third column). The superior ability of the second- and third-year group is clearly shown, but it is left to the first-year group to reveal the greatest improvement from first to second rating: a fact which suggests that changes are due at least in some degree to changes in taste or in ability on the part of the raters. Had they been due wholly to the effects of a second contact with the poems the effects would surely have been more marked with the more able and experienced readers (provided of course there was room for such improvement).

The science forms stand out for their tendency to prefer the false poems (though they show considerable improvement), and the boys are remarkable for lack of change. The division by sexes yields in fact two groups which show comparatively little improvement in the third column of the table: division of substantially the same sample into other groupings shows more change, particularly the division by maturity. This may suggest that the contributions of the two sexes in some way complement each other. (Boys' order correlated with girls', 0·37: first-year group with second- and third-year group, 0·74.)

It may be added that a similar relationship in this respect seems to hold among individual experts making up the group of eighteen. When correlated in turn with the true/false order the highest coefficient gained by any expert (and by one only) is 0·91. The average order of all experts nevertheless achieves the same result, 0·91: suggesting that their individual ratings pool to give something more like a composite order than a simple average.

(viii) *The convergence of preference*

It is satisfactory to have found a convergence of preference upon the true poems (though with exceptions), but the question may be asked, what caused this convergence? Does the change indicate an improvement in individual poetic taste during 6 months? Or the fact that a second attempt at a poem, after an interval, takes a reader a stage nearer to a true appreciation? One shred of evidence was offered in §(vii) above, but obviously the evidence of the present experiment, in which the *same* poems were presented twice, cannot provide conclusive answer, and it might be more prudent to suggest that the two explanations need not be regarded as entirely separate. If repeated readings of a poem over an interval of time could be shown to lead to a truer appreciation of that poem, it would have to be demonstrated that such an experience did not also tend to induce a more favourable approach to *other* poems before it could be claimed that the two questions were, in normal circumstances, independent.

Time spent reading poetry seems to be the core of the matter. Changes of opinion with regard to poetry on the part of people who never read poetry (for what the subject is worth) could clearly have no connexion with the training effect of repeated readings. The training given to the raters taking part in this experiment would be so diverse that no generalization can be made about it except to say that in all but a minority of cases some time would be spent by the candidates in reading and considering poetry. (The science forms, who would be likely to spend less time reading poetry, certainly showed a poorer result.) In general the results suggest that, provided the habit of reading poetry was sufficiently present for some poems at least to make an impact, the study of these poems, repeated after an interval, was likely to give rise to changes of opinion in the right direction.

It is suggested that the best teachers of poetry have always worked upon the assumption that time spent in reading good poetry will, in spite of initial setbacks, contribute towards the formation of good poetic judgement.

V. SUMMARY AND CONCLUSIONS

A battery of eight short poems by modern writers and seven counterfeit 'poems', prepared for the occasion, was presented to 221 raters, eighteen of whom were accounted experts. The eight true poems represented four categories based upon two bipolar factors extracted in an experiment by Eysenck (1940): namely, simple and complex poetry; and poetry of restrained and abandoned feelings. Raters arranged the poems in the order of their preference and gave comments in explanation. One hundred and twenty of the original raters carried out the process a second time after an interval of 4–6 months.

Preference orders were converted into scores roughly representing the frequencies of normal distribution. Results were analysed by use of Student's test as applied by Fisher (1936) to the means of small samples: and by the method of analysis of variance.

The experts were found to show a significant preference for true poems (particularly in the categories 'abandoned' and 'restrained') and significant preferences for restrained poems and either for complex poems (in a majority of cases) or for simple poems (in a minority of cases). The whole sample of 221 raters showed no discrimination whatever in favour of true poems as a class, but were significantly influenced by all four of the bipolar categories. Raters were divided into groups in various ways and results further analysed.

The difference between simple and complex poetry was found to be an issue in poetic preference with all groups of raters. Notably, the less experienced preferred the simple poems, the more experienced the complex poems, though the distinction seemed less influential with the experts themselves. The distinction between abandoned and restrained poetry was found to operate one-sidedly, in favour of restraint, a result which may well reflect the revolt in recent years against romanticism in poetry.

Categories derived by analysis of poetic judgements, categories of actual opinion, are obviously of very great importance in the study and in the teaching of poetry. It is perhaps easy to make the mistake of going further and regarding them as in some way prescriptive—a role for which they would be even less fitted than the categories laid down from time to time by experienced critics. For the teacher they will be a starting-point, rather than any kind of destination.

It was inferred from ratings and comments by raters that the more experienced reader tends to respond to a poem either by committing himself fairly positively (both in endorsement and in rejection) or by suspending judgement upon it. It is the sign of the reader least experienced in poetry that he rejects what he cannot comprehend simply on the grounds that he cannot comprehend it.

The primary obstacle to poetic judgement was found to lie in the intervention, between the reader and the expression of the poet, of preconceived ideas, attitudes and sentiments. It is suggested that stock responses and other forms of prejudice may sometimes be adopted out of instinctive caution: a protective measure, so to speak, against the new experience where it is of an emotional kind; and a refuge instead in the familiar emotional situation preconceived.

The re-test revealed a significant increase in true-poem totals as a whole, that is a convergence of preference upon the true poems. The primary condition for its occurrence would seem to be that poetry should be read and returned to.

There seemed some evidence that average orders of preference ratings might in some cases be regarded as 'composite orders', a pooling of complementary opinions. As an observer's angle of view may be unique and his view yet complementary to other views of the same object from other angles, so greater objectivity might sometimes be gained by pooling poetic judgements than would appear possible by any consideration of the overlap of individual views.

The writer gratefully acknowledges the help of Prof. P. E. Vernon with this research.

REFERENCES

BRITTON, J. N. (1953). A factor analysis of poetic preferences.*
ELIOT, T. S. (1934). *Elizabethan Essays.* London: Faber.
EYSENCK, H. J. (1940). Some factors in the appreciation of poetry, and their relation to temperamental qualities. *Character Pers.* IX, 160–7.
FISHER, R. A. (1936). *Statistical Methods for Research Workers*, 6 ed. Edinburgh: Oliver and Boyd.
RICHARDS, I. A. (1935). *Practical Criticism.* London: Kegan Paul.

APPENDIX

Details of test items

(Counterfeit poems are referred to in brackets)

No.
1 (Emblem)
2 The Youth with the Red Gold Hair—Edith Sitwell
3 (Word out of Season)
4 (The Sea)
5 Bright Clouds—Edward Thomas
6 Francesca—Ezra Pound
7 (Faithful unto Death)
8 (If there is a no-man's land)
9 The Villagers and Death—Robert Graves
10 I dreamed that I stood—W. B. Yeats
11 I tell you the generations of man—Archibald Macleish
12 In Memoriam, Easter 1915—Edward Thomas
13 (Autumn)
14 (Casualty List)
15 September Fires—Herbert Read

Examples of Counterfeit Poems

(3) Word out of Season

Yours was always the voice I heard
However many spoke beside.
Amidst the clattering tongues one word
From you would bring me to your side.

When you were with me all you said
I treasured in my heart, until,
Goodbye being over, in my head
Your precious words were ringing still.

I took your words to bed with me:
Now nothing else is mine to keep—
The voice is gone—the mocking words
Destroy my peace, deny me sleep.

---

* Copy may be had on loan from the Technical Information and Documents Unit, Department of Scientific and Industrial Research.

(13)

Autumn

Autumn has draped the boughs in gold and gossamer;
The woods are hers: her mists are silver in the air,
And on the mist, what bitter scent of burning years!
What requiem for rose and daffodil, what tears
For leafy bowers soon naked and sunny petals shed!
We walk the dusk and speak of summer's glories dead.

And yet the woods are lovely still! No requiem
Need steal their ageing beauty, nor disfigure them
With bitter thoughts. Silver glides the swollen stream
Beneath the branchy elms: the moulting tree-tops seem
As lovely in the autumn light as ever dappled sun and shade
That corruscated leaf and limb upon the grassy glade.

Love too from spring to autumn moves, and when the sun
That shone on golden lad and lass at last has run
Its fiercest noon-day flight, and first approaching night
Spreads lengthened shadows on the path, the mellow light
Of true affection, flooding as the sun's late rays,
Shall shine in tears and laughter down love's autumnal days.

(14)

Casualty-list

Proud Pharoah, long from flesh, preserves his dust
In golden pomp: no sacrilegious touch
Disturbs his dreams; a thousand years he must
Alone have waited on the perfumed flesh.
Pale silken queens their lonely vigil keep
Bedecked for flattery that none shall pay;
Only the bone remembers: none could weep
For love of lovely dust or moving clay.

Whom the gods love die swift: no burial
Bestows their bones with runes and lingering grief,
'And some there be have no memorial',
So quick and clean is death. O fate be brief!
All flesh is grass and fiercely burn the fires of war:
No rites of grief shall stay them at the opening door!

(*Manuscript received* 1 *December* 1953)

A SEX DIFFERENCE IN PREFERENCES FOR SHAPES

By W. A. McELROY

Psychology Department, University of Sydney

I. *Introduction* (pp. 209–211). II. *Test administration and test results* (pp. 211–213). III. *Interpretation of the sex differences* (pp. 213–216). IV. *Summary* (p. 216). *References* (p. 216).

I. INTRODUCTION

A feature of many experiments concerned with visual aesthetic preferences has been the variety of sources from which the subjects have reported aesthetic enjoyment. In addition to other elements, Valentine (1950) has suggested that 'the general preference for curves over straight lines may be something very fundamental and connected with the appeal of curves in the human form'. For our everyday aesthetic preferences, Thouless (1947) has stated a possible corollary of psychoanalytic theory in the following terms:

> It cannot be doubted that the pleasure taken in the appearance of the human form is the result of the appeal that it makes to the tendency with a more powerful content of pleasure than any other—the sex instinct. The form most strongly appreciated as beautiful is almost invariably that of the opposite sex....Those who appreciate the widespread symbolism of sex will not doubt that sexual associations lie at the bottom of the beauty perceived in many other objects. At least, we must recognize that much of the pleasure in works of art may be due to their unrecognized appeal to the sex emotions.

From biological principles, it is stated here that a basic sex difference in aesthetic appreciation is to be found in the male preference for the female form, and the female preference for the male form. Considering the role played by symbolism in mental life, it is alleged that a modified but derived sex difference must also exist with respect to suitable symbolic representations of male and female forms. This possible corollary suggests, therefore, that perception by the male of any suitable symbolic representation of the female form may be accompanied in consciousness by a surplus of pleasurable affect denied to the female, with a converse process applying to perception by the female of any suitable representation of the male form.

The female form differs essentially from the male in its curved aspects, especially with respect to the full development of the breasts, and it appears likely that this quality of the female form may be suitably represented symbolically by simple curved diagrams. In addition, the psychoanalysts have maintained that pointed forms are typical male symbols (Freud, 1911).

Considering these aspects of the symbolism theory, and its possible corollary, a twelve-item aesthetics test was constructed. This test is exactly reproduced in Fig. 1. Each item consists of a pair of line drawings, one labelled '*A*' and the other '*B*'. For each item, one of the drawings was selected on account of its curved features and the other was selected because of its possible symbolic phallic properties. For each of the twelve items shown in Fig. 1 an asterisk has been placed against the label of the line drawing intended to represent disguised aspects of the female form. (These asterisks were not present in the test proper). Six items, irregularly distributed throughout the test, had curved drawings first; the other six items showed curved drawings last as the *B* number.

That the shapes chosen for the test have in some way either 'masculine' or 'feminine' significance is supported in part by the work of other investigators[1] and their premises. In experiments reported by Franck (1946) round forms are regarded as feminine designs, while masculine designs consisted of elongated forms. Alschuler & Hattwick (1947), in observation of the aesthetic expression of children, have reported that straight lines or elongated figures are characteristic products of aggressive moods of the more 'masculine' children. They also observe that children with negative feelings towards their mothers tended to destroy rounded forms in their drawings. Supporting evidence also comes from Krout (1950), who found that her subjects responded to a semicircle as a 'female' form, and responded to a straight line as a 'male' form.

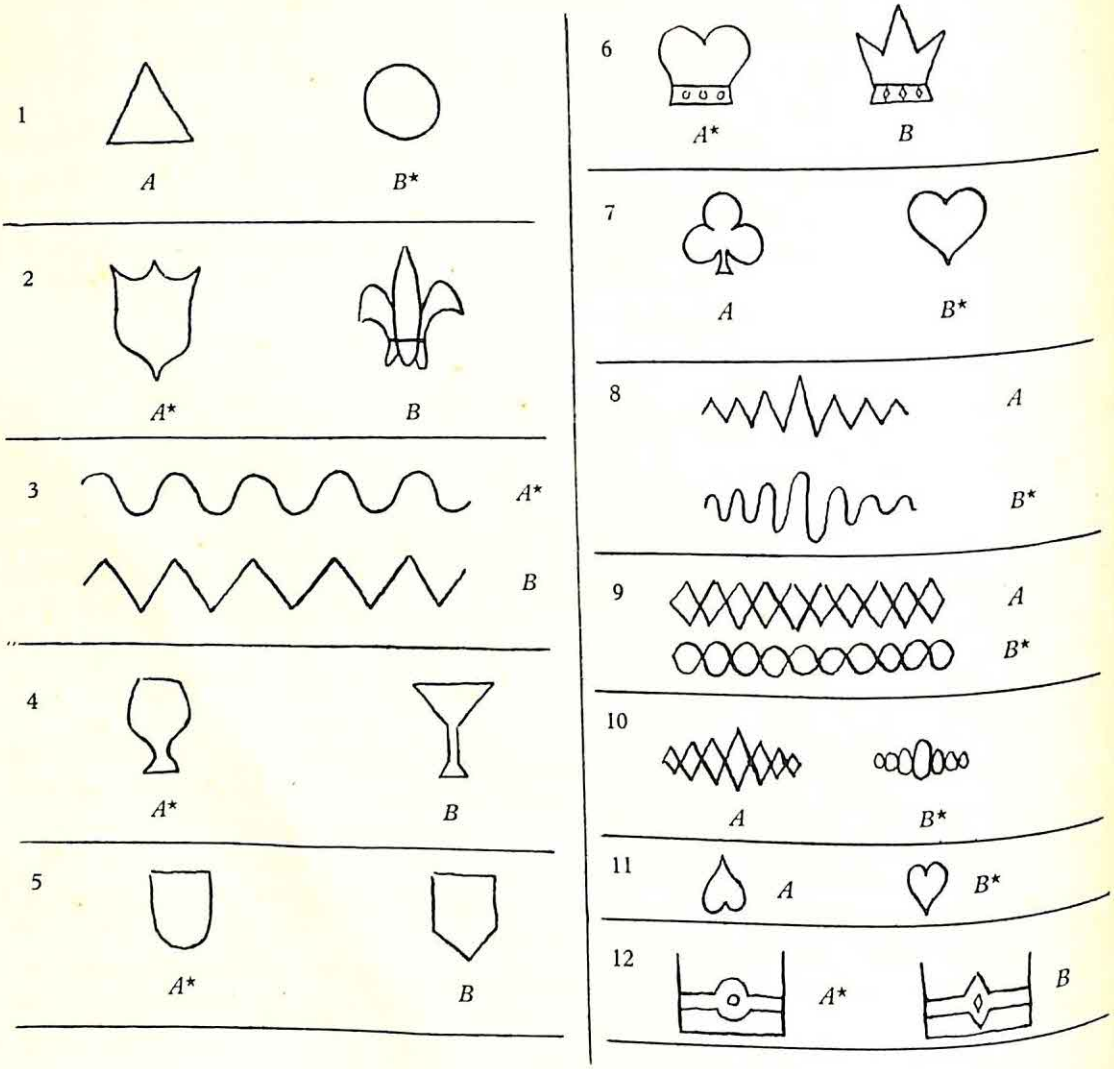


Fig 1.

---

[1] Item 1 of the aesthetic appreciation test reported here consists of a triangle and a circle. This pair of stimuli were used by Dr Florence Kluckhohn in a study of phantasy content in relation to pattern of culture (personal communication). It was found that the subjects—U.S. students—tended towards a phantasy stereotype of describing the triangle and circle as 'a man and his girl friend'. An interesting feature of the study was the tendency of the subjects towards personalization of figures.

Let us suppose that a number of subjects of both sexes are instructed to state on record sheets individual preferences for *A* or *B* for each of the twelve items of the test. It is possible to predict from the corollary previously stated that sex differences will exist in the responses of the group, on account of the postulated 'surplus of affect' being distributed in opposite directions for the male and female groups. It is to be noticed that this corollary does not justify the prediction that the number of responses other than male preferences for 'female' shapes and female preferences for 'male' shapes will not differ significantly from zero, since many other factors such as the symmetry and evenness of the drawings must be expected to influence test results. It should also be noted that identification responses, i.e. Lipps's (1905) empathic responses, might also be a mechanism of preference. If this were the primary basis for preference, a sex difference might be expected to exist in the opposite direction to the prediction stated above, so that a greater proportion of female than male subjects would record preferences for 'female' shapes.

II. Test administration and test results

It was considered undesirable to apply the test to an experimental group of adult subjects, particularly University students. It seemed probable that the content of the test would be more disguised for younger subjects. The actual experimental group consisted of 779 children from three different Scottish schools, comprising 380 boys and 399 girls between the ages of 9 and 16. Three hundred and seventeen subjects (148 boys and 169 girls) were between the ages of 9 and 12 and formed the older half of a Scottish Elementary school; 33 subjects (17 boys and 16 girls) were aged 11–13 and formed a preparatory class about to enter a Scottish Secondary school; and 429 subjects (215 boys and 214 girls) were between the ages of 11 and 16 and formed the first four years of a Scottish Senior Secondary school. (The schools were non-feepaying schools in the suburbs of an industrial city.)

The test items were duplicated upon sheets headed 'Picture Test'. These sheets were distributed to the subjects, who were tested in their classrooms in groups of approximately forty subjects. In most cases each group contained members of both sexes. Blank paper was distributed for the records of preferences, and the subjects were asked to place name and age upon this paper, in addition to marking the numbers one to twelve in the margin. The subjects were then informed that they were to try a 'Picture Test', and were asked to look at test item number one (Triangle and Circle). The instructions given to all groups were as follows:

> Look at number one on the sheet with the pictures. Opposite this number are two pictures—a triangle and a circle. The triangle is called *A* and the circle is called *B*. What you have to do is to decide which you like better. If you like the triangle better than the circle, then put *A* opposite number one on the sheet with your name on it. If you like the circle better than the triangle, then put *B* opposite number one on your paper. When you have done this, continue with the other items in the same way. Does anyone not understand what they have to do?...Remember that in this test there are no right or wrong answers. We want to know what pictures *you* like best.

During the test, children were discouraged from discussing items or answers with one another. Answer sheets were collected from individual subjects as soon as they had been completed. The frequencies of the responses to the twelve test items for 380 male subjects and 399 female subjects are shown in Table 1. Also shown are the proportions of the male and female groups recording preferences for those diagrams intended as disguised 'female' configurations, and for those intended as disguised 'male' configurations. An inspection of these data for each item shows that either (*a*) a greater proportion of male than female subjects recorded preferences for 'female' configurations, or (*b*) a greater proportion of female than male subjects recorded preferences for 'male' configurations. The disparity in aesthetic judgements between the groups could have been produced by either possibilities or both.

Table 1

	Males		Females		Proportions of preferences						
					'Female' shapes		'Male' shapes				
Item	*A*	*B*	*A*	*B*	Male	Female	Male	Female	χ^2	c	r_p
1	297	83	355	44	0·218	0·110	0·782	0·890	12·08	3·48	0·15
2	88	292	54	345	0·232	0·135	0·768	0·865	9·72	3·12	0·12
3	210	170	155	244	0·553	0·388	0·447	0·612	21·36	4·62	0·16
4	214	166	204	195	0·563	0·511	0·437	0·489	1·95	1·31	0·05
5	148	232	78	321	0·389	0·195	0·611	0·805	32·45	5·29	0·21
6	212	168	155	244	0·558	0·388	0·442	0·612	22·71	4·41	0·17
7	206	174	264	135	0·458	0·338	0·542	0·662	12·08	3·48	0·12
8	241	139	309	90	0·366	0·226	0·634	0·774	17·57	4·19	0·15
9	281	99	350	49	0·261	0·123	0·739	0·877	18·18	4·27	0·18
10	215	165	265	134	0·434	0·336	0·566	0·664	8·42	2·90	0·08
11	119	261	163	236	0·687	0·591	0·313	0·409	6·11	2·47	0·09
12	190	190	162	237	0·500	0·406	0·500	0·594	7·22	2·69	0·09

$\chi^2 = 169{\cdot}85$; $r_p = 0{\cdot}14$; $c_{mf} = 13{\cdot}03$.

For each item the significance of this disparity was first estimated by applying a χ^2 test of significance between the frequency of the responses which appeared to be in agreement with the psychoanalytic corollary (i.e. the frequency of male preferences for 'female' shapes plus female preferences for 'male' shapes) and the frequency of the responses which appeared to be in disagreement (i.e. the frequency of male preferences for 'male' shapes plus female preferences for 'female' shapes). In each case the square root of the resultant χ^2 was extracted to give an exact measure of the critical ratio of the difference between the proportions of the sex groups recording preferences for 'female' configurations, or between the proportions of the sex groups recording preferences for 'male' configurations. The critical ratio c gives the magnitude of the tested disparity in terms of a normal deviation with unit standard measure, and the value for each item is entered in Table 1. An inspection of these values shows that the sex difference is fully significant at the 0·01 level for all items except nos. 4 and 11, although the latter item provides a sex difference significant at the 0·05 level. It will be seen that pooling the results for all items gives a resultant χ^2 of 169·85. The square root of this value provides an estimate of the critical ratio of the difference between the proportions of the sex groups recording preferences for 'female' shapes—or between the proportions of the sex groups recording preferences for 'male' shapes—over the experiment as a whole, and this critical ratio c_{mf} has the value 13·03. The level of significance is therefore extremely high, and the deviation is in the direction predicted from the possible corollary of psychoanalytic theory. Considering the direction of the deviation, it is apparent that identification responses have not provided the principal basis for preference, although it is likely that they have contributed towards the total experimental results.

From an analysis of the data given in the table, the mean male test 'score' is 5·22 preferences for disguised 'female' configurations or 6·78 preferences for disguised 'male' configurations. The estimated standard deviation of the male 'scores' is 1·71, and the standard error of the mean is 0·493. The mean female test 'score' is 3·75 preferences for disguised 'female' configurations or 8·25 preferences for disguised 'male' configurations. The estimated standard deviation of the female 'scores' is 1·60, and the standard error of the mean is 0·462.

It will be seen that although the difference between the sex groups tends to be fully significant for most items, the correlation between sex and type of preference is not itself large. A fourfold point correlation r_p is given for each item in Table 1, and this varies from 0·05 for item no. 4 to 0·21 for item no. 5. Over the twelve items the average correlation is 0·14.

To examine the relationship between age and preference for shapes,[1] a simple dichotomous classification of the test data was made according to whether the subjects had stated an age level below 12 years, or had stated age as 12 years or above. The number of preferences for 'female' shapes is now given for these two age groups, and the critical ratio of the difference between the older and younger divisions of the sex groups with respect to the number of preferences for 'female' shapes is also given (as estimated from the root of a χ^2 test of significance):

	Male		Female	
	No. subjects	No. preferences	No. subjects	No. preferences
Below 12 years	161	745	182	739
12 years or above	219	1238	217	757
	$c_d = 4{\cdot}16$		$c_d = -2{\cdot}96$	

From these data it will be seen that a significant rise in male preferences for 'female' configurations has occurred between the younger and older male groups. A corresponding smaller but significant decrease in female preferences for these configurations has also been recorded.

III. INTERPRETATIONS OF THE SEX DIFFERENCES

It was stated earlier that it is possible to entertain some alternative hypotheses other than psychoanalytic or biodynamic ones which might account for some aspects of the results obtained. These hypotheses will be considered first.

(*a*) An inspection of the test reproduced in Fig. 1 shows that the drawings varied in their accuracy. Since rectilinear drawings are easier to draw than curvilinear ones, it is possible that the curvilinear drawings contain a greater proportion of inaccuracies than the rectilinear ones. It can be further maintained that girls are more critical of minor inaccuracies than boys, hence their preferences. A close inspection of Fig. 1 suggests that this hypothesis might certainly apply to test items nos. 1 and 4, where the circle[1] and the wine glass with the curved profile can be seen to contain more inaccuracies than their corresponding 'male' shapes. Yet it should be noted that these items, where the 'female' shapes contain probably greatest inaccuracies relative to the 'male', are not the ones which produce the greatest deviations in test results. In fact item no. 4, with the 'female' shape perhaps poorest in accuracy relative to the 'male' shape, provided the smallest deviation of all the items. We note, further, that some of the intended 'male' configurations involved in some cases the drawings of as much curvilinear outline as the corresponding 'female' configurations, viz. items nos. 2, 7 and 11. Yet these items provided wide deviations in test results. In items nos. 3, 5, 6, 8 and 12, the inaccuracies do not appear to be confined to any one type of diagram rather than the other type, whereas item no. 9 suggests that the rectilinear outline is, if anything, more poorly drawn than the curvilinear one. Whether this hypothesis could account for the relationship between age and preference also appears doubtful. Nevertheless, it is possible that the factors suggested by the hypothesis may have contributed towards test results as a whole.

---

[1] In a paper entitled 'Experimental Tests of some Aspects of Psychoanalytic Hypotheses' read (by Dr R. W. Pickford in the absence of the present writer) at the Annual General Meeting of the British Psychological Society in April 1951, it was stated that age appeared to make little difference to the reported sex difference in preferences for shapes for the age levels studied. (In that paper the number of male preferences for 'female' shapes was added to the number of female preferences for 'male' shapes and classed as responses in 'Agreement' with psychoanalytic hypotheses. Other responses were classed in the 'Disagreement' category). This statement was based upon two empirical observations. First, evidence of the sex difference was found at all the age levels studied. Secondly, a comparison between the extremes of the age groups had revealed apparent consistency in the magnitude of the sex difference as given by the ratio of Agreement to Disagreement responses. It is clear from the more appropriate analysis of the test results given above, that a relationship with age does exist.

(*b*) Some of the shapes are likely to have specific associations for the subjects. For example, with item no. 2 (*A*, shield, *B*, fleur de lys) it is possible that *A* might suggest to the boys a shield or coat-of arms, whereas outline *B* might be more familiar to girls as a pattern found on materials for curtains, etc. In this way the associations of the boys and the girls might be different for the same shape. Prof. Sir Cyril Burt asked several subjects the reasons for their preferences after these had been obtained. Some of the associations are given below:

> *Item No.* 2. The majority (of the boys) undoubtedly preferred the shield—usually because it was a *shield*, i.e. a symbol of warlike adventure....Two of the girls said that the fleur de lys reminded them of Joan of Arc, and added that was why they liked it. Only one referred to the shield as a *shield*, and she added that that was why she *disliked* it.
>
> *Item No.* 3. Several girls said they would have preferred the wavy line if it had been better drawn. Two of those who preferred the zigzag line said it would be much easier as a pattern for stitching.
>
> *Item No.* 4. The boys liked the idea of drinking wine out of the big rounded wine-glass. The girls tended to prefer the other wine glass, which two of them called a 'champagne glass'; and they obviously thought champagne glasses were much more interesting than the old Georgian pattern.

It is certainly true that some features of the cultural background which foster adventurous boys and less adventurous girls may have contributed towards the test results obtained for item no. 2 in the way the above associations suggest. Yet whether this would account for all of the deviation obtained with item no. 2 seems open to question. Other features of the cultural background might suggest that boys should not prefer the shield. For example, the fact that the fleur de lys is shown on the Boy Scout's badge could be used as evidence of this. Whether similar special features of the cultural background other than those associated directly with sex (as with the drinking glasses suggested by the associations for item no. 4) could account for all of the deviations for all of the items, especially the less meaningful ones, is again open to doubt.

In considering reasons given by the subjects for their preferences, the question arises as to how far it is possible to advance *valid* reasons for aesthetic preferences, and also how far reasons advanced are sole determinants of the preferences. No doubt it is at times possible to advance valid reasons for aesthetic preferences, but for many aesthetic judgements this would not necessarily be the case. Likes and dislikes, for example, that are based upon conditioned emotional reactions, are unlikely to be accompanied by conscious awareness of the true reasons for aversions and preferences. It is possible that associations of the kind reported by the subjects studied by Prof. Burt have contributed towards test results. Yet the possibility of over-determination of object choice might also be considered. Thus if any boys tended to select the shield through operation of some unconscious determinant they might be even more inclined to select the shield since it is a *shield*—a justifiable choice of boys.

---

[1] Several female subjects studied by Prof. Burt commented on the fact that the circle was not *perfectly* circular, and added that if it had been they would probably have preferred it.

(*c*) Since the sex difference was predicted from a suggested corollary of psychoanalytic theory, it is at once apparent that the test results can be interpreted as supporting the psychoanalytic symbolism theory. By this theory, as stated by Jones (1918), a powerful affective and conative process, i.e. some behaviour tendency invested with libido, which is held in a state of psychical repression, and in consequence is denied entry to consciousness and is not permitted free external expression, may be symbolized by some idea in consciousness, owing to the existence of a process of unconscious identification. In this way the occurrence of a conscious idea or percept can be a partial means of gratification of a repressed idea, and hence such an occurrence may be accompanied by an affective state derived not as the result of the occurrence of the conscious idea or percept as such, but as the result of the partial expression of the unconscious repressed idea. It is clear, however, that the experiment cannot be regarded as forming a complete experimental analogue of the symbolism theory, since the content of the symbolism appears to be less extensive than the theory as stated by Freud. It certainly appears unlikely that *all* of the types of objects that Freud (1911) believed to be typical female symbols which occur in dreams (e.g. stoves, rooms, and wardrobes) would be perceived by the male in accompaniment with a surplus of aesthetic enjoyment; or that tree-stumps, umbrellas, or nail-files would necessarily provide additional aesthetic pleasure to the percepts of females.[1] It should be noted, however, that the sex difference could have been predicted from certain well-established principles of learning, without involving psychoanalytic concepts.

(*d*) The well-established principle of 'stimulus generalization' could have been used as the basis of prediction of sex differences in responses to the test items. Miller (1935) has reported generalization of fear responses in animals from one food cup to a highly similar drinking cup. Similarly, Masserman (1946) has described generalized fear of 'all food boxes and associated configurations' after cats had been placed in a conflict situation centred around one particular food box. From this principle, with reference to 'associated configurations of the male and female form', we could have predicted sex differences in preferences. If this principle were used as an explanatory one, then it might be possible to account for the significant rise in male preferences for female-associated configurations through a tendency for the primary stimulus to be reinforced with greater pleasurable affect.[2] For the female group, however, a similar explanatory principle might appear unlikely. It would be possible to account for the relationship with age in terms of generalization of fear responses, since it is around 12 years and later that problems of female adolescent development are likely to occur. These possible explanations of the relationship between age and type of preference are, of course, speculative. It is further possible that any valence that the shapes had for the children could have been assimilated unconsciously from the valences placed upon different shapes by adults, and that the relationship with age might reflect the increase in assimilation of post-adolescent valences.

---

[1] Evidence from clinical psychologists does suggest, however, that many unusual objects can provide aesthetic enjoyment for individual persons. A. A. Brill (1914), for example, quotes an interesting case of a lady who collected coloured plates of mushrooms, and suggests that this has symbolic significance.

[2] In psychoanalytic terms, this relationship with age could be explained as occurring through an increase in repressed affect.

Whatever interpretation is accepted, it is possible that the sex difference is culturally delimited.

IV. SUMMARY

This paper reports the existence of a sex difference in preferences for shapes as shown by the results of an aesthetics test carried out with Scottish schoolchildren. The sex difference was predicted from a possible corollary of psychoanalytic theory, so that it is possible to consider the test results as supporting Freudian principles. Some alternative hypotheses, however, can be advanced to account for the results obtained. It is further possible that the sex difference cannot be explained in terms of any one hypotheses considered alone. Whatever interpretation is accepted, the writer is at present concerned only incidentally with describing the construction of the aesthetics test and method of prediction of the sex difference, and is primarily concerned with reporting the test results.

The writer is indebted to Prof. Sir Cyril Burt for many detailed comments upon the empirical data reported in this paper. The writer has called freely upon these comments in preparation of the text. He is also indebted to Prof. Burt for supplying the reasons given for aesthetic preferences made by some subjects studied by him using the test described here.

REFERENCES

ALSCHULER, R. H. & HATTWICK, L. W. (1947). *Painting and Personality.* Chicago: University Chicago Press.

BRILL, A. A. (1914). *Psychoanalysis*, p. 234. Philadelphia: Saunders.

FRANCK, K. (1946). Preferences for sex symbols and their personality correlates. *Genet. Psychol. Monogr.* XXXIII, 73.

FREUD, S. (1911). *Die Traumdeutung*, 3rd ed., p. 211. Leipzig and Wien: Deuticke.

JONES, E. (1918). *Papers on Psychoanalysis.* Chapter: 'The theory of symbolism.' London: Baillière, Tindall and Cox.

KROUT, J. (1950). Symbol elaboration test. *Psychol. Monogr.* no. 310, LX, no. 4.

LIPPS, T. (1905). *Psychologische Studien.* Leipzig.

MASSERMAN, J. (1946). *Principles of Dynamic Psychiatry*, p. 109. Philadelphia: Saunders.

MILLER, N. E. (1935). A reply to 'sign gestalt or conditioned reflex?' *Psychol. Rev.* XLII, 80.

THOULESS, R. H. (1947). *General and Social Psychology*, 2nd ed., p. 452. London: University Tutorial Press.

VALENTINE, C. W. (1950). *Psychology and Its Bearing on Education*, p. 439. London: Methuen.

(*Manuscript received* 17 *November* 1953)

A NEW TACHISTOSCOPE AND CYCLING TIMER

By W. C. R. WITHERS

Psychology Department, Institute of Psychiatry, University of London

I. *Tachistoscope* (pp. 217–218). II. *Timer* (pp. 218–219). III. *Summary* (p. 219). *References* (p. 219).

I. TACHISTOSCOPE

The need for an improved tachistoscope and timer arose out of an 'apparent movement' study conducted by Brenner (1953) in this department. In this study a modified Dodge tachistoscope similar to the one described by Merryman & Allen (1953), was used. In operation the first figure was seen directly through an eyepiece and half-silvered mirror; the second figure being seen by reflexion in the half-silvered mirror. This necessitated complete illumination of the cards from one side, so that not only the stimulus figure was illuminated, but also the inside of the box. Further, if the pattern on the cards was at all complex, correct alinement of both cards became extremely difficult to achieve and the apparent movement occurred in two directions simultaneously.

To overcome these disadvantages, a direct viewing tachistoscope (Fig. 1) was devised.

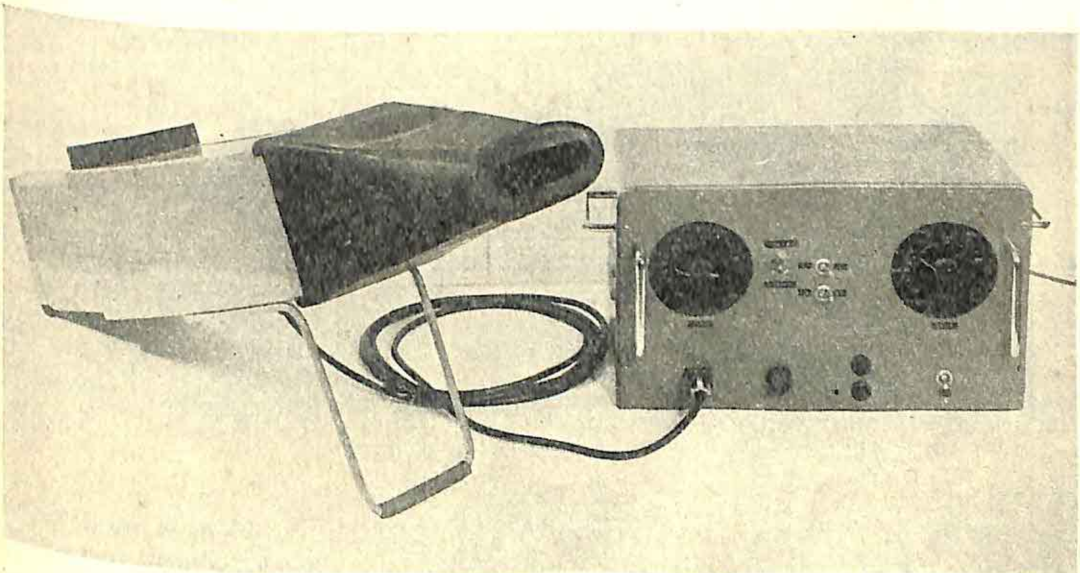


Fig. 1. New tachistoscope and electronic timer.

This tachistoscope consists of a 6 in. square metal tube 15 in. in length, painted black inside; a rubber viewing mask at one end, 11 in. long (ex Radar) fits closely round the subject's eyes. The other end fits over a square sectionalized lamp housing containing four standard 220 V. 5 W. neon lamps, each mounted to illuminate a quarter of the box. A slit in the body of the tachistoscope in front of and above the lamp housing, allows cards of black bakelite to be inserted. Stimulus figures are cut out of these cards. A metal mask in front of the cards holds them firmly against the lamp housing, so that light can neither be seen round the edges of the cards, nor escape from one section of the lamp housing to another. The bakelite cards themselves have white typing paper pasted on the reverse side to give even illumination of the figures. The top two lamps are on at all times, and the stationary part of the figure is cut out in front of them. To obtain apparent movement, the lower two lamps are switched on in turn by an electronic timer. To allow the lamps to be connected in different combinations, a tag panel is mounted on the back of the lamp housing. The apparatus is fitted on a stand to present the eyepiece at eye level to the seated viewer.

This tachistoscope has the advantage that only the figures are seen, and there are no alinement difficulties.

Other applications

By altering the lamp connexions, other experiments can be carried out. For example, flicker-fusion tests using one lamp only: or apparent rotary motion can be obtained by adding a further metal section between the top two lamps.

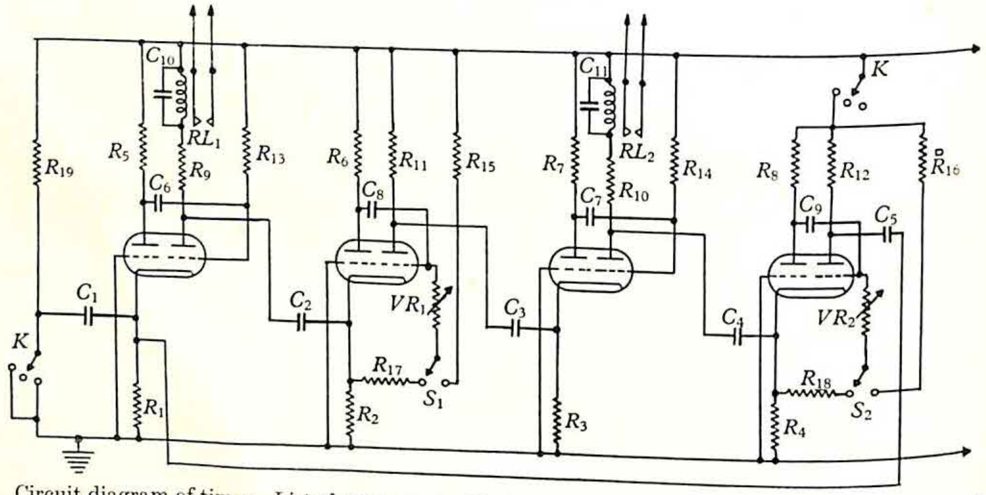


Fig. 2. Circuit diagram of timer. *List of components.* Resistors: R_1, R_2, R_3, R_4, 2·4 kΩ ± 10%; R_5, R_6, R_7, R_8, 100kΩ; R_9, R_{10}, 6·8 kΩ; R_{11}, R_{12}, 10 kΩ; R_{13}, R_{14}, 1·5 MΩ; R_{15}, R_{16}, 390 kΩ; R_{17}, R_{18}, 330 kΩ; R_{19}, 10 MΩ; VR_1, VR_2, 5 MΩ (linear). Capacitors: C_1, C_2, C_3, C_4, C_5, 0·001 μF.; C_6, C_7, 0·1 μF.; C_8, C_9, 0·15 μF. Relays: RL_1, RL_2, 3·4 kΩ, Siemens high-speed relays. Switches: S_1, S_2, single pole 2-way (range switches); K, double pole 3-way key switch. Valves: All 6 SN7, each half matched within 15%.

II. TIMER

The timer for the first tachistoscope was an electronic one in which a multivibrator circuit drove round a uniselector which connected the lamps for the required period. This was a simple arrangement, but had the disadvantage of producing clicks giving auditory cues to the subject; also this timer had no facilities for cycling the stimulus figure.

A new timer was developed to overcome these defects, the circuit diagram of which is presented in (Fig. 2).

Operation

When the experimenter presses down the operate key (K) the first 'flip flop' (2) turns on the first of the two lower lamps for a fixed period of 50 msec. (controlled by $R_{13}C_6$). On flopping back, a sharp negative pulse is injected into the cathode of the second 'flip flop', flipping it over for a period determined by the setting of (VR_1). When this 'flip flop' reverts it operates the third 'flip flop' and the second lamp for 50 msec. ($R_{14}C_7$).

For cycling, the operate key is set in the upward position, connecting a further stage, the fourth 'flip flop'. This causes the whole cycle of operations to be repeated after an interval determined by the setting of VR_2. With the key (K) in the centre position, the timer is inoperative.

Other applications

Several applications are possible. For example, a similar timer has been arranged by the writer to give a conditioned and overlapping unconditioned stimulus, for use in conditioned reflex experiments. The overlap is achieved by connecting C_2 from the second 'flip flop' directly to the operate key (K), the extent of the overlap then being determined by VR_1.

Circuit notes

The relays for the neon lamps are Siemens high-speed relays, and are mounted on sponge rubber in a box of acoustic boarding. With the timer in its case, no operating sounds can be detected at about 3 ft. away.

The capacitors C_{10} and C_{11}, shown in the circuit diagram (Fig. 2), are to prevent high oscillatory voltages from being generated and passed to the next stage.

The coupling from anode to cathode of succeeding 'flip flops' (which the writer believes to be an original arrangement) obviated the need for phase-inverting stages.

The initiating network ($R_{19}C_1$) is arranged so as to produce only negative pulses.

The duration of the flashes is fixed at 50 msec. The intervals between them are variable from 18 msec. to 3 sec. (calibrated with a 'Chronotron' Timer).

Power supplies are: h.t. 150 V. neon stabilized, and 220 V. d.c. for the neon lamps obtained directly from the mains via a metal rectifier and smoothing capacitor.

III. SUMMARY

The construction and working of a new tachistoscope and new electronic timer are described. Directly viewed cut-out figures are used in this tachistoscope. The timer is a silent, electronic model giving either cyclic or single exposures.

The writer is indebted to May Brenner and W. D. Furneaux, of this department, for their many helpful suggestions relating to the design of the tachistoscope and timer.

REFERENCES

BRENNER, MAY (1953). Unpublished Ph.D. Thesis, University of London.

DICKINSON, C. J. (1950). *Electrophysiological Technique*, p. 43. London: Electronic Engineering.

MERRYMAN, J. G. & ALLEN, H. E., (1953). An improved electronic tachistoscope. *Amer. J. Psychol.* LXVI, 110.

(*Manuscript received* 11 *September* 1953)

CORRESPONDENCE

THE EDITOR
British Journal of Psychology

Dear Sir,

In the current volume of the *Annual Review of Psychology* there is a paragraph by Masserman, Gross & Pechtel (1954) commenting on my article on experimental neuroses that recently appeared in this *Journal* (Wolpe, 1952*a*). The paragraph not only ignores all the major conclusions of the article, but in addition contains several inaccurate and misleading statements to which attention is drawn below.

The paragraph opens with this sentence: 'Wolpe reported his observations on 12 cats subjected to grid shocks with or without learned feeding responses, and in general confirmed the results of similar experiments by previous workers.' This statement glosses over the crucial point that *because* neuroses were produced by shocks *without* the animal having previously learned feeding responses in the experimental situation, motivational conflict in Masserman's sense could not have been essential to their production. Masserman (1943) and other workers (e.g. Dimmick, Ludlow & Whiteman, 1939) had earlier assumed that conflict was essential, because feeding responses were involved in all their experiments, and controls without feeding were not done. Since my experiments contradicted this assumption there is no justification for the suggestion of unreasonableness on my part conveyed in Masserman, Gross & Pechtel's second sentence—'Nevertheless he objects to Masserman's concepts of motivational conflict...'; for such an objection is precisely what was substantiated.

The same sentence continues with an assertion that I propose that 'experimental neuroses are induced (*a*) by "ambivalent stimulations" arising from difficult discriminations, delay before reinforcement, or rapid reversals, and (*b*) "noxious stimuli" which interrupt learned responses'. Anyone would gather from this that items (*a*) and (*b*) comprise my *explanation* of experimental neuroses, whereas, in fact, they are merely some of the many experimental situations I enumerated as productive of neuroses. (Noxious stimuli are effective *whether or not* they interrupt feeding responses.) Actually, I tried to show that all the various effective situations are anxiety-evoking and that experimental neuroses are due to the conditioning of powerful anxiety responses to previously 'indifferent' stimuli.

In the light of the foregoing it is necessary to reject also Masserman, Gross & Pechtel's final remark that my formulations 'eventuate in about the same procedures both for inducing and treating the neurotic deviations'. When new facts change theories, new procedures eventuate. In the present context this has already happened. Both experimentally (Wolpe, 1952*a*) and clinically (Wolpe, 1952*b*, 1954) treatment procedures have emerged that are very different from those generated by Masserman's theory.

Yours faithfully,
J. WOLPE

Department of Psychiatry
University of the Witwatersrand

References

Dimmick, F. L., Ludlow, N. & Whiteman, A. (1939). A study of 'experimental neurosis' in cats. *J. Comp. Psychol.* xxviii, 39–43.

Masserman, J. H. (1943). *Behaviour and Neurosis*. Chicago: University of Chicago Press.

Masserman, J. H., Gross, Z. & Pechtel, C. (1954). Abnormalities of behavior. *Ann. Rev. Psychol.* v, 263–80.

Wolpe, J. (1952*a*). Experimental neuroses as learned behaviour. *Brit. J. Psychol.* xliii, 243–68.

Wolpe, J. (1952*b*). Objective psychotherapy of the neuroses. *S. Afr. Med. J.* xxvi, 825–9.

Wolpe, J. (1954). Reciprocal inhibition as the main basis of psychotherapeutic effects. (*Arch. Neur. and Psych.* To be published.)

The Editor
British Journal of Psychology

Dear Sir,

The publications listed below[1] will show that during the past fifteen years my colleagues and I have noted and specifically controlled the phenomena reported by Dr Wolpe. I shall, therefore, not further dilute the literature by replying in detail to Dr Wolpe's comparisons of our data and interpretations, including our quite relevant experiments on the production of 'masochistic' cats that avidly *seek* electroshocks. Instead, it may be of greater interest to comment on a general tendency among some psychologists (and, alas, not a few psychiatrists) to pre-empt a few words in the English language, take out a special patent as to their meaning and then attempt to live off the royalties. An example is Dr Wolpe's use of 'motivational conflict', an expression which he apparently wishes to restrict to a clash between 'learned feeding responses' and grid shocks. Actually, rational reflection must expand such terms to encompass any disruptive impasse between (*a*) any *dynamic configuration of accustomed adaptive patterns* (including, in a cat's universe, the opportunity to explore an experimental cage without being rudely electroshocked), and (*b*) any new set of experiences that runs counter to such expectations—and is thereby rendered operationally 'noxious' or 'traumatic'. Were this not true, human neuroses would quite literally be confined to people who had either been caught in an electrical short-circuit or struck by lightning at the moment they were pushing an accustomed button to secure food from an automat.

---

[1] Publications with references to related work by the author and his associates are:

Neurosis and alcohol. *Amer. J. Psychiat.* ci, 389. (1944.)

Dynamic psychology of war-time communications and morale. *Dis. Nerv. System*, v, 101. (1944.)

Principles of Dynamic Psychiatry; including an Integrative Approach to Abnormal and Clinical Psychology; with a Glossary of Psychiatric Terms. Philadelphia: W. B. Saunders. (1946.)

Effects of cerebral electroshock on experimental neuroses in cats. *Amer. J. Psychiat.* civ, 92–9. (With Mary Grier Jacques.) (1947.)

Psychological medicine and world affairs. In Harris, Noel Gordon, ed.: *Modern Trends in Psychological Medicine.* New York: Hoeber. (1948.)

A biodynamic psychoanalytic approach to the problems of feeling and emotion. In Reymert, Martin L., ed.: *Feelings and Emotions. Second Mooseheart Symposium*, ch. iv, pp. 40–75. (1950.)

Effects of direct interrupted electroshock on experimental neuroses. *J. Nerv. Ment. Dis.* cxii, 384–92. (With Alex Arieff, Curtis Pechtel and Harold Klehr.) (1950.)

Experimental approaches to psychodynamic problems. *J. Mt Sinai Hosp.* xix, 639–52. (1953.)

Psychoanalysis and biodynamics—an integration. *Int. J. Psycho-Anal.* (Suppl.), xxxiv, 13–43. (1953.)

Role of olfaction in normal and neurotic behavior in animals. *Psychosom. Med.* xv, 396–404. (1953.)

Faith and delusion in psychotherapy. *Amer. J. Psychiat.* cx, 324–33. (1953).

Conflict-engendered neurotic and psychotic behavior in monkeys. *J. Nerv. Ment. Dis.* (In the Press.)

Neuroses in monkeys: a preliminary report on experimental observations. *Annals N.Y. Acad. Sci.* lvi (no. 2), 253–65. (With C. Pechtel.) (1953.)

As to Dr Wolpe's therapeutic inferences from his work: I regret that his recent publications and manuscripts are not as yet available here, so that I must be content to await further information from him with keen interest. Parenthetically, I do not quite understand what he means by 'treatment procedures...very different from those generated by Masserman's theory', since, in my *Principles of Dynamic Psychiatry*, based in part on our research, I tried to encompass all principal therapeutic methods from Ra to Rank. However, if Dr Wolpe will accept an exchange of stakes very likely in my favour, I shall wager a set of autographed copies of the publications listed above against correspondingly inscribed reprints of his forthcoming papers (*a*) that the therapeutic principles he derives from his studies are already widely known and applied, and (*b*) that they were probably also in use in the Asklepiad Sanatoria of the ancient Hellenes. And, as an instance of my trust in Dr Wolpe's scientific probity: if he will but consult the appropriate historical, experimental and clinical sources, he can be the sole judge of the matter.

I shall be glad to discuss the subject further with Dr Wolpe through more usual channels, but appreciate this opportunity to communicate and clarify our views.

Yours faithfully,
JULES H. MASSERMAN

The Medical School
North Western University
Chicago
Illinois

CRITICAL NOTICE

By SIR GODFREY THOMSON

Uppsala Symposium on Psychological Factor Analysis, 17–19 March 1953. Nordisk Psychologi's Monograph 3. Distributed by Almquist and Wicksell Stockholm.

This three-day meeting of psychologists and statisticians was planned in order to further the understanding of their respective approaches to factor analysis. The remit was to discuss 'what psychology needs and what statistics can provide'. There were forty-four participants, including thirty-two from Sweden, five from Denmark, one from Norway and one from Finland. Of the other five, one each were from Princeton and Bombay, and three from the United Kingdom. These last three (Peel, M. S. Bartlett and Lawley) played a major part in the proceedings as the writers of invited addresses to the first three sessions. The other main papers were by Wold and Whittle of Uppsala and Rasch of Copenhagen. A paper by Lundberg of Lund was also briefly dealt with but is not available until it appears in *Psychometrika*. All the other papers are here printed except Whittle's which had already appeared in the Skandinavisk Aktuarietidskrift for 1952. The discussions are also reported.

It was a very important move in the development of factor analysis to bring some British statisticians who have shown interest in factor analysis into fertilizing contact with the famous Swedish School. Others will single out other names, but I was particularly impressed by reading that Cramér of Stockholm was present and took the chair at one session.

Peel's paper was a readable and illuminating review of past and present, not in the main mathematical. He suggested recommendations for the future, and the Symposium report ends with a valuable list of ten such, decided upon after general discussion.

Bartlett's paper was a very clear account of the mathematical history of factor analysis, and in his last three pages he adds another original contribution to the many he has made, namely a discussion of an attempt to transcend the hypothesis of linearity in the fundamental equation of factor analysis.

Lawley (the pioneer in the application of the method of maximum likelihood to factor analysis) gave an exposition of his new Method III, and Whittle's paper was another approach by maximum likelihood. Wold gave arithmetical illustrations of Whittle's method used on artificial test-scores.

The quite important papers of Rasch on simultaneous factor analysis in different populations, and of Lund on a ball and wire model, nevertheless stood somewhat apart from the main issue, which was on the papers by Lawley and Whittle.

Whittle's paper presents I think the most novel idea. He succeeded in estimating in one stage both the loadings of the factors in the tests, and the factor values possessed by the people tested. He does it by fitting, not the matrix of covariances or correlations, but the matrix of scores, an oblong matrix. He makes no assumption about the distribution of factor values in the population, but only about the ratios of the residual variances in each test. Lawley assumes, in addition, the factor values to be normally distributed. Their results are very similar but differ by a factor, which each confessed he was as yet unable to explain.

The ideas underlying Whittle's attack, as he says in his opening paragraphs (and indeed Lawley's too), go back to an article by Gale Young in *Psychometrika*, February 1941, and, I would add, to an earlier paper in that journal for September 1936 by Eckart and Gale Young, who discussed the problem of fitting most closely (least squares) one matrix by another of lower rank. In the case of a correlation matrix this is done by expressing the full matrix as a triple matrix product, with a diagonal matrix of its latent roots in the middle, flanked on both sides by the matrix of latent vectors. The best fitting matrix of a given lower rank is then obtained by replacing the necessary number of latent roots, beginning with the smallest, by zeros, and then multiplying the three matrices together.

This process gives new values not only to the correlations but also to the diagonal cells, and it interests me very much to observe that these values in the diagonal cells are larger than the communalities which would give the same lower rank. The 'specifics', that is, are smaller, and I have often protested against their very large variances produced by the method of communalities. The sum of squares of the discrepancies of the Eckart and Young approximation, *taken over all the cells including the diagonal*, is much less than that of the communality matrix, if there also the diagonal discrepancies are included.

So far I have spoken of a square matrix, but (and here is the link with Whittle) Eckart and Young also show how to apply these principles to an oblong matrix (such as a matrix of scores). Such an oblong matrix too can be expressed as a triple product, with a diagonal matrix in the middle (with however some extra columns of zeros). The entries in this diagonal matrix are the square roots of the latent roots of the covariance matrix made from the oblong matrix; and the pre-multiplying matrix contains the orthogonalized latent vectors of that covariance matrix. The post-multiplying matrix is found by solving for it in the matrix equation, and turns out to be a horizontal slab of an orthogonal matrix. If the dimensions of the oblong matrix were n tests and N persons, the dimensions of the three matrices are

$$(n \times n)\,(n \times N)\,(N \times N) = (n \times N),$$

but the last $N-n$ columns of the middle matrix are all zeros, and the last $N-n$ rows of the post-multiplying matrix are not required, and need not be found. The best-fitting matrix of lower rank r can then be found by replacing all except the r largest entries in the middle matrix by zeros and multiplying the three matrices together.

Five years later Young (after referring to Lawley's Method I) applied this plan, using maximum likelihood, to test-score matrices, assuming that the errors are normally distributed but that the matrix of the reciprocals of the standard deviations of the errors is of rank one. He knew of no solution to the more general case. It appears to be here that Whittle's paper fits in, and Lawley's Methods II and III. Whittle describes Method II as 'an appropriate extension of Young's procedure'; but it broke down, and Method III takes its place.

Young speaks of 'errors' where Lawley speaks of 'specifics' and Whittle of 'residuals', all meaning the same defect from full variance in the diagonal cells. Lawley at the outset of his Uppsala paper speaks of these differences, for they are more than merely verbal. There are, he said, two models in use. In the first (Whittle's) both the loadings and the factors are treated as parameters to be estimated rather than as variates. The residuals are assumed to be independent random errors normally distributed.

In the second model (Lawley's) these residuals represent not only error but also specific factors. If the same assumption of normal distribution is made for these specific factors it seems reasonable to make it about the common factors, which are therefore to be treated as independent normal variates, leaving only the loadings as parameters to be estimated. Wold, in discussion, called Lawley's analysis a special case of Whittle's, but Lawley did not agree. 'There is surely', he said, 'a deeprooted distinction between the conceptions of the factor values as statistical variates and as systematic parameters.'

About the fact that these two methods differ by a factor, Bartlett said that this may be analogous to the difference between his own method of factor estimation and Thomson's (i.e. mine). Bartlett's minimize over the specifics, mine ensure that the mean square deviations of the factor estimates are a minimum over all persons in a group having the same set of scores. In both applications the size of the specific factors has already been decided in order to reduce the rank of the matrix. It would be interesting to make some arithmetical comparisons, helped by Wold's paper.

These methods of Lawley and of Whittle resemble closely a principal components analysis. When the residual or specific variances are all equal they reduce exactly to a principal component analysis. It is likely that in orthodox factor analysis Lawley's Method III will be preferred to Whittle's method, because it reproduces the correlations better. All depends of course on whether we should give pride of place to this, or to reproducing the whole matrix, including the diagonal, by one of lower rank.

Lawley's Method I is perhaps still more likely to be preferred by factor analysts. It works perfectly well in practice except only in cases where 'impossible' matrices crop up, which, as Bartlett said at Uppsala, does not invalidate the method. Lawley has not yet been able to give a fully rigorous proof that it leads to 'efficient' statistics, but it seems almost certain that it does so.

The Uppsala Symposium is a milestone in the progress of factor analysis. What I would like to happen next is an elementary seminar with explanations as simple as possible of the points discussed at such a high level at Uppsala. The members would have to be given prerequisite reading, and examined on it!

PUBLICATIONS RECENTLY RECEIVED

Selection for Industrial Leadership. By A. G. Arbous. London: Oxford University Press. 1954. Pp. xvi+180. 30*s*.

In 1948 the South African National Institute for Personnel Research was called on to assist in selecting a small number of administrative trainees, that is potential high-grade executives, for a large industrial corporation. A procedure very similar to that of our own Civil Service Selection Boards (though somewhat less thorough, for reasons outside the Institute's control) was applied to 219 candidates, of whom 16 were selected. Since, however, more than two-thirds of the candidates were already in the corporation's employment, it was possible to follow up a population of 116 rejected as well as the 16 selected candidates, and to obtain comparable ratings a year later on their potentialities as administrators, thus by-passing the very difficult problem of correcting follow-up results for selectivity. An adaptation of the Otis Group Test yielded a validity of 0·53, Matrices 0·43, assessments of a written project (similar to the CSSB 'Island Story') 0·53, and assessments of group discussion and committee sessions 0·60. The multiple correlation based on these reached 0·68. However, this probably represents somewhat poorer validity than that obtained in CSSB selection, since the variance of the candidates in *g* and education was obviously unusually large. Though several other interesting techniques were employed—a general knowledge test, an employment interview by each candidate with a 'stooge', a clinical interview by a psychologist, and a sociometric questionnaire—their possible contributions to the procedure are not indicated. A subsidiary study in another organization suggested that the corresponding validity of conventional selection by application blank, testimonials and Board interview, was only of the order of 0·20.

The value of the book lies less in the psychological results that we have outlined than in its review of the literature of group methods, and its contributions to the statistical methodology of selection. The former provides an excellent and critical survey of WOSB, CSSB and OSS work, though exception must be taken to one quite crucial point—namely the applicability of the formula for indirect selectivity. Arbous argues that in CSSB selection, for example, the follow-up group was selected on a subjective and unspecifiable combination of variables. True, the Final Selection Board mark was derived from such a complex combination, but the selection—so far as the statistician is concerned—was based exclusively on the FSB mark, and therefore satisfies the conditions for applying this supremely useful technique.

Nearly half the volume deals with the interpretation of validity coefficients which, as Arbous makes clear, have little meaning apart from the practical conditions of application of the selection procedure. Everything depends on the proportion of candidates it is proposed to select, and on the training wastage and wastage of wrongly rejected candidates that can be permitted. Techniques, and specimen graphs, are given so that employers can determine the precise effects of putting the procedure into practice (but though the description is lucid, an employer would need a good deal of statistical training to follow it). The reviewer is not entirely convinced of the value of such techniques, since validity coefficients (particularly multiple ones) vary so much from one sample to another. Moreover, the very common situation where the selector expects an improvement in the quality of selectees by raising the pass-mark (e.g. all educational selection) is not covered. Nevertheless, Arbous's approach is undoubtely stimulating, and any psychologist concerned with selection problems would profit from studying his contributions.

P.E.V.

The Geometry of Mental Measurement. By Sir Godfrey H. Thomson. London: The University of London Press Ltd. 1954. Pp. 60. 6*s*. 6*d*.

This little book is based on three lectures given by Sir Godfrey Thomson in London University in January 1953. His aim is to construct a geometrical model in terms of which the main formulae used in the statistical analysis of test results can be interpreted. Any psychologist who has a slight mathematical background and is 'willing to think spatially' should certainly read it, for the author's lucid style and omission of any irrelevant details will help many readers to appreciate in a new way the significance of techniques with which they are already familiar.

The correlation between two tests is envisaged as the cosine of the angle between two axes (representing the tests); starting from this basic idea the author leads on to a discussion of partial correlation, regression and factor analysis, in terms of angles and projection in two or more dimensions. By confining himself to the extraction of one factor, with only three tests, he shows clearly the geometrical meanings of 'communalities' and 'factor loadings', and the problems involved in the estimation of factors. The author does not concern himself here with sampling problems, calculations, rotation of axes, or the psychological interpretation of the relationships discussed; his treatment is essentially abstract and, for the geometrically inclined, it is very refreshing.

M. A. CREASY

Rituals of Rebellion in South-east Africa. By MAX GLUCKMAN. Manchester: Manchester University Press. 1954. Pp. 36. 3*s.* 6*d.*

Social tensions, it is now fashionable to hold, are bad; much writing and investigation is directed towards the discovery of reducing or eliminating tensions in society. This assumption that human society would work better without tension, conflict and frustration is rejected by most social anthropologists; in their analysis of social systems they assert that cohesion in a society is achieved by the mutual opposition of parts and by the pressure of external enemies; 'a society is knit together by its inner cleavages'. Among most primitive peoples men are agreed about the right form of society and struggle for power within it; ambition leads them to become rebels, not revolutionaries, and usurpers, not heretics, so that every rebellion is a reaffirmation of the rightness of the existing order. The latent and enduring rivalries and conflicts also find expression at annual ceremonies, when subjects insult and taunt their rulers, princes challenge the king, and women pantomime the roles of men.

In this Frazer Lecture for 1952 Prof. Gluckman examines certain agricultural ceremonies in which social conflicts are vividly dramatized, and contrasts them with those of our own society in which 'royalty is blatantly extolled, the virtues of the nation are magnified and the country is glorified'. He argues that 'every social system is a field of tension, full of ambivalence, of co-operation and contrasting struggle', and that where the common moral order is not questioned, men emphasize by acting out the conflicts either directly or symbolically, the social cohesion within which the conflicts exist. Where the social order is itself questioned, as in Western society, political ceremonies cannot be used for catharsis in this way.

This short work is scattered with suggestions about the connexion between hostility and social cohesion, and should be read by all who maintain that tension and conflict are always socially destructive.

J. A. BARNES

The African Mind in Health and Disease. A Study in Ethnopsychiatry. By J. C. CAROTHERS. World Health Organization, Geneva. 1953. Pp. 177. 10*s.*

It is a pity that an excellent review of the psychology and psychiatry of the African should have been spoiled by premature generalization. The literature of the subject is well covered—though the anthropologists might have been given more space in view of the importance of the argument from culture that occupies much of the book. There is certainly no other work which so well expounds what we know of the African. But how little we do know is clear from the short bibliography of only 191 items, covering not only the psychology and psychiatry of the African, but also his nutrition, culture, physical anthropology and ethnology. The paucity of material should, perhaps, have warned the author of the danger of generalization.

His generalization is that the African in Africa south of the Sahara is more like his fellows in that enormous continent than a European is like other Europeans. We are told that he is similar both in culture and personality, and that the latter is mainly the product of the former. It may be that both these generalizations are true, but if they are then they must be counted as very lucky hunches—or penetrating insights—for the evidence the author advances is not at all convincing.

For example, he says that children are much desired by African parents, are indulged up to the time of weaning and that from birth they live in a smooth, warm rhythmic world not far removed from the prenatal one. No authorities are given for this statement, which is meant to apply to the whole of southern Africa and to all individuals. But a few sentences below this (p. 43) it is asserted that there is another side to the picture; that young children are often left in charge of a small sister quite unfitted for the responsibility, and that they may be forced against their will to eat unsuitable foods. The author implies that the indulgence is of major importance whilst the frustrations which may be produced by the young sister are unimportant. But he gives no evidence of either the amount or relative importance of the two kinds of treatment, and neither does he tell us in what precise ways the sister may frustrate the young child.

All discussion of the cultural environment is of this kind. Carothers merely asserts the cultural conditions to be of a certain kind, and, what is more dangerous, asserts that they have specific psychological effects. Weaning, for example, he says will be very traumatic to the African child. This assertion is taken from the study of Ritchie, which itself is based largely on impressionistic data and was conducted in a very small area of Africa. Your reviewer has just concluded a detailed study of sixteen African children and has found that the dire conditions asserted by Ritchie to result from weaning do not appear. By the seventh week after weaning the child appears little affected by the process and, indeed, there is some evidence that his development towards maturity has been aided. Also, it is clear that at eighteen months the African child is almost ready to wean himself.

The generalizations about the psychology of the African are based on the work of seventeen authors. Much of this work is very restricted both in numbers and scope, not all of it is empirical and some is frankly anecdotal. To generalize about the African from such data is dangerous, and becomes more evidently so when it is realized that most work has been done on intelligence, and in this exact field the results are doubtful. If it is difficult to establish differences by means of well-developed techniques, how much more difficult it will be to be certain of results obtained with less exact methods.

The generalization that the African suffers from a 'fragility of the higher psychic functions...' may be a good guess but it is unwarranted to attribute this disturbance to the frontal lobes when we know so little about frontal lobe function.

This book is perhaps best regarded as providing us with an excellent review of the psychology of the African together with yet another stereotype. The stereotype presented does not, after all, differ markedly from the ordinary one of the African as a shiftless, lazy and happy-go-lucky soul. It may be that he is such a person, but a great deal more evidence than we have is necessary to establish this as a fact.

R. C. ALBINO

The Integration of Behaviour, Vols. I and II. By THOMAS M. FRENCH. Chicago: University of Chicago Press; London: Cambridge University Press. 1954. Pp. xii+272. 37*s*. 6*d*. net (vol. I). Pp. xii+367. 49*s*. net (vol. II).

Anyone undertaking to write a series of five volumes intended to elucidate the integration of behaviour must at least be allowed a sense of adventure. Dr French is well endowed with this sense. This is clear from the two volumes of the series which have so far appeared.

His basic assumption is this: the student of behaviour must first analyse it into its supposed elementary reaction patterns before attempting to demonstrate how these are combined. Dr French endeavours to identify these elementary patterns by a detailed study of a patient (not his own) during two years of analysis. His declared aim is not to prove anything but the more modest one of illustrating a method.

The first volume has been given the subtitle 'Basic Postulates'. The choice of this subtitle remains a mystery for it covers a wide miscellany of topics, including studies of frustration, the learning process in therapy, and the influence of sleep. The second volume explores the integrative process in dreams. Here we must pay tribute to Dr French for his courage in coming to grips with a problem of outstanding importance, though the academic psychologist does not recognize it as such, in spite of the fact that he dreams away at least a third of his life. Dr French here assumes a twofold task: first to relate the different possible meanings of a dream to one another and to show how they fit together in a closely-knit logical structure, and second to reveal how different dreams of the same person are connected. He undoubtedly succeeds in making many illuminating comparisons between the patient's dreams dreamed on different occasions and in varying circumstances.

Dr French draws inspiration from many sources, chiefly from Freud but also from Lewin, Tolman and even Hull. He employs a Hullian type of symbolism the esoteric character of which is likely to restrict its appeal, and his conception of 'integrative capacity', i.e. the amount of influence on behaviour of a particular insight at a particular time, is conceived on analogy with Hull's notion of habit strength.

Dr French would have employed his clinical gifts to better purpose if he had confined them to a more limited goal. He seems to be out of his theoretical depth. The reader would be willing to sacrifice a little integration of behaviour in exchange for some integration of theory. The fifty chapters in the first volume and the fifty-nine chapters in the second are often notes and unintegrated fragments rather than chapters. The thinking which should have been done by the author is left to the reader.

In the three volumes yet to come Dr French proposes to discuss variations in the patient's behaviour from time to time. To prejudge these future volumes would be unjust. For the author's sake it is to be hoped that they will be more coherent than their predecessors.

JOHN COHEN

Introduction to General Psychology. By E. J. ASHER, JOSEPH TIFFIN and F. B. KNIGHT. Boston: D. C. Heath and Co.; London: G. G. Harrap and Co. Ltd. 1953. Pp. xvi + 515. 25*s.*

The appearance of a new text-book on General Psychology is always an interesting event. In this volume the authors have presented a balanced survey of contemporary psychology, arranging the various subtopics (chapters) in a logical order. Emphasis is placed upon psychology as an empirical science. While ample factual material is presented, even in this introductory treatment the authors very rightly find space to raise some of the controversial issues in contemporary psychology. A certain amount of over-simplification of complex issues is inevitable. Whilst the reviewer would not himself rely on any single existing text for the beginner, he would add the present volume to the limited list of those which might well be adopted for such introductory courses. Amongst those responsible in the U.K. for initiating the beginner into the mysteries of modern psychology some may consider the book too long, but it is difficult to see how the same amount of material could be condensed further. Like many texts in use, this naturally reflects the American 'culture pattern'. The chapter on 'Abilities and Aptitudes' places considerable emphasis on apparatus tests which will seem to some of us to belong to the 1930's. However, the chapter is rounded off with a useful and critical discussion of the interpretation of test results. The diagrams and other illustrations in the book are simple, clear and informative, and the general layout is excellent. By modern standards the price of twenty-five shillings certainly represents good value for money, a point which will recommend the book to the student, as one which he might purchase.

ROLAND HARPER

Human Behaviour: Psychology as a Bio-Social Science. By LAWRENCE E. COLE. New York: World Book Co.; London: G. G. Harrap and Co. Ltd. 1953. Pp. xii + 884 37*s.* 6*d.* net.

Developmental Psychology. By ELIZABETH B. HURLOCK. New York, Toronto and London: McGraw-Hill Book Co. Inc. 1953. Pp. x + 556. 48*s.*

Essentially, *Human Behaviour* is an expanded and more up-to-date version of the author's deservedly popular *General Psychology*. Despite its title, it contains a good deal of information about animal experiments, and though it is thin on physiological psychology it is otherwise notably comprehensive. The book has no particular theoretical standpoint, the 'bio-social' approach being confined mainly to the rather laboured paragraphs which open some of the chapters. There is a certain lack of perspective—inevitable, perhaps, when so wide a field is covered by a single writer—but this is a minor defect in a book which most readers will treat mainly as a work of reference.

One of the most valuable features of *General Psychology* is the clear and balanced account of Freudian theory which is concealed in the chapter on 'Thinking'. This account has been expanded in the final section of *Human Behaviour*, but part of the new material is noticeably inferior in quality: the content is slight, and the clear and workmanlike style of the rest of the book degenerates into 'colourful' journalese in which, for example, Janet's exhaustion concept is said to be, in Freud's opinion, 'as far from the harsh realities that go on within the mental apparatus as a prettified chromo from the dear, dead past'. However, these unfortunate pages, with their dragged-in references to Pareto and Niebuhr and Kierkegaard, form only a small part of a long, meaty, lucid and thoroughly competent textbook.

Developmental Psychology has little truck with depth-psychology. It is a full and systematic account of the changes, physical and mental, involved in growth and senescence, and it contains some excellent advice, at the commonsense level, on how best to maintain adjustment at the various stages ('adjustment' is the author's favourite word). Dr Hurlock is not afraid of the obvious, e.g. 'Divorce is an indication of poor marital adjustment....Legal separations and desertion are likewise indications of poor marital adjustment'; but she is conspicuously lucid and sensible, and her book, though described by the publishers as a second-year text-book, should be a valuable source of information and advice to the layman.

MARGARET KNIGHT

Psychoanalytic Theories of Personality. By GERALD S. BLUM. New York, Toronto and London: McGraw-Hill Book Co. Inc. 1953. Pp. xviii+219. $3.75, 32*s*.

This is an extremely learned summary of the theories of Freud, and of practically all those of his followers as have diverged from him in any way. It is arranged in the order of the individual's development from pre-natal life to adulthood, and is liberally illustrated with the results of experimental work, beginning with the increase in foetal heart-beat and movement in response to sharp, loud noises, and going up to such studies as an attempt to relate oral traits in adults, as measured by a self-rating scale, to age of weaning. It is not illustrated at all with the case-material which is the principal source of the psycho-analyst's knowledge. A beginner attempting to use it as a short-cut to such knowledge, or even as a guide to further reading, would gain little understanding, and would be bewildered by a book in which the profound discoveries of Jung and Melanie Klein are treated on a level with scores of minor variations of Freudian theory propounded by relatively unimportant people. For anyone, however, who has a fair amount of psychoanalytic insight, it will be valuable as a refresher course, a reference book, and a stimulus to thought and further discovery.

JANE DARROCH

Hypnotism: an Objective Study in Suggestibility. By A. M. WEITZENHOFFER. New York: John Wiley and Sons, Inc.; London: Chapman and Hall Ltd. 1953. Pp. xvi+380. 48*s*.

This book is the first real attempt to present all the available experimental evidence in this field since C. L. Hull published his classic *Hypnosis and Suggestibility: An Experimental Approach*. In view of the mass of material that has accumulated since, the attempt was certainly well worth making, and it must be judged to have been, on the whole, remarkably successful. The author, aided by his training in the exact sciences, has a good eye for well-controlled experimentation, a critical mind, is well read in the subject, and obviously has personal experience in the field.

The documentation is very thorough, the bibliography running to over 500 titles. Of major omissions which could alter the author's conclusions I have only been able to find one. In his discussion of hypnosis and dissociation on page 99, Weitzenhoffer quotes the work of Mitchell, which is negative, but not that of Nagge, which is positive, and being derived from a larger number of cases appears more trustworthy. This is an important point as Weitzenhoffer bases some of his theoretical conclusions on this study.

Welcome as the detailed discussion of the experimental literature is, the most important contribution Weitzenhoffer has to offer is probably his theory of the hypnotic process. He eschews monistic theories and puts forward the view that 'Hypnosis is the outcome of not one but of several temporally overlapping and interacting processes'. The two main ones the author considers to be those of *homoaction* and *generalization*; in the later stages of hypnosis a third process, namely a narrowing of consciousness, is introduced. The theory is ingenious and has the usual advantages and disadvantages of theories involving the interaction of several processes, none of which can be specified or measured with any degree of accuracy. It would probably be possible to account for most hypnotic phenomena along these lines, but it would be very difficult to make predictions which could invalidate this theory. In spite of this criticism, theoretical formulations in the field of hypnotic phenomena have generally been so obviously untenable that Weitzenhoffer's formulation stands out as a much more reasonable interpretation of the very complex phenomena under discussion than any alternative theory known to the reviewer.

H. J. EYSENCK

Trauma, Growth and Personality. By PHYLLIS GREENACRE. The International Psycho-Analytical Library and the Institute of Psycho-Analysis. The Hogarth Press 1953. Pp. xi+308. 30*s*. Ed. by Ernest Jones.

This volume is a selection of papers made during the past decade and published in various journals. They have a special interest because the author, Dr Greenacre, is a clinician through and through, which means that what she writes, even when it could be better written and even when it is without form, is never without interest. Fortunately Dr Greenacre has wide clinical experience which includes paediatrics and especially the care of infants and of newborn infants. She also has experience of adult psychiatry from the mental hospital angle and to this she has added the psychoanalytic training and many years of full-blooded psychoanalytic work. Moreover, in the latter she has undertaken treatments of borderline cases where the matter of management cannot be thought of as outside the analytic work since the patient of whatever age is at times an infant in the transference situation. If the patient is being a new-born infant or even an infant acting out birth memories Dr Greenacre knows where she is because of her exceptionally wide experience apart from that gained in doing classical psychoanalysis.

The first chapter on The Biological Economy of Birth (1945) interests me most, because it is unlikely that there are more than a very few properly trained and fully experienced analysts who are competent to refer to the phenomena of actual birth. One starts with the stimulus provided by Freud's suggestion that the phenomena of anxiety are related to the phenomena of the birth experience. Freud however was not (at any rate in the early days of psychoanalysis) thinking of the individual's memories but of a build-up of the birth experiences of a million years in the race of human beings. The difficulty is to know how much of the human psyche is present at the time of birth. One can describe the phenomena of the birth trauma but many of these are present even in the case of an anencephalic monster. Appearances therefore are deceptive. At the other extreme from anencephaly there are the adult's fantasies about birth which are notoriously coloured by knowledge gained at later dates, so that whenever birth experiences are remembered they are liable to be mixed up with pseudo-memories and even with clear visual impressions which cannot possibly have been based on the sight memories of the birth era. Complex fantasies of birth, however, do not exclude the possibility of birth memories, and what is needed if progress is to be made is knowledge of the psyche of the human infant at the time of birth so that the meaning of the traumata to the infant can be assessed. I think there is no study of this subject at all comparable to that which appears in the first chapter of this book and also in the next on The Predisposition to Anxiety. No attempt is made to reach a final conclusion but the subject is opened up and the author is capable of making sane and objective judgements through being an observer of infants at the same time as being a psycho-analyst with a grip on psychoanalytic theory. Room is left for the conception of a normal birth and for a specific anxiety pattern that may be associated with birth by Caesarian section.

It would seem to me that Dr Greenacre could enrich her work by bringing to her study of the problem the conception of environmental adaptation as something which at the beginning belongs to the infant. If the infant is at first not differentiated from the environment, then environmental adaptation, if good enough, is part of the infant, although as we see it this adaptation is the work of a mother who at the beginning loves in physical terms. In describing the management of patients who seem to be dealing with birth trauma in the transference situation Dr Greenacre describes her own technique but she does not actually join this up in theory with the primary narcissism of the infant, explaining that in her management she is acting a part of the individual infant before individuation has become a fact.

In further chapters Dr Greenacre's wide experience continues to enrich her studies of the early stages of the development of female sexuality with a special emphasis on the effect of anatomical structure on the super-ego formation which is in some respects different in girls from what it is in boys. Also there is a valuable comment on the effect of assault on girls in the pre-puberty age.

For the clinician this is an exceptionally valuable addition to the International Psycho-Analytical Library. Psychoanalytic theory is basically accepted and yet is opened up for modification and development. This comes from the author's clinical experience in which she brings together in her technique for the analysis of borderline cases a true appreciation of classical psychoanalysis and first-hand experience in the management problems of early infancy and of birth.

D. W. WINNICOTT

Psychological Reflections. By C. G. JUNG. Collected and Edited by Jolande Jacobi. London: Routledge and Kegan Paul. 1953. Pp. xxviii + 342. 30*s*. net.

This anthology, collected and edited by Dr Jolande Jacobi, was first published in Zurich in 1945, the year of Jung's seventieth birthday. Dr Jacobi set herself the considerable task of presenting to the general reader a panoramic view of Jung's contribution to modern thought.

The book contains a selection of over a thousand quotations. They are mostly statements of a general nature that can stand by themselves. Within a framework consisting of sections and subsections, the quotations are arranged, not in chronological order, but in meaningful sequences designed to illustrate particular themes in the most appropriate way.

In this book Jung the philosopher takes precedence over the clinician, and the subject-matter of the quotations deals with wide spheres of human life and activity. The nature of the psyche, primordial images, dreams, men and women, problems of self-realization, the life of the Spirit, Fate, Death and Renewal—all these and many others.

This is not a book that can be read easily or quickly, but it has been most thoughtfully compiled. For those who can digest general conclusions without demanding the particular evidence for them, this book succeeds in conveying the nature and quality of Jung's contribution to modern thought more vividly than any summary of his work could possibly do.

Many things have been said, and will continue to be said, for and against Jung. He is, *par excellence*, a controversial figure; inspiring to some and infuriating to others. But, as Dr Jacobi remarks in her preface, 'No-one who does not stand aside from real life will be able to avoid coming to terms with Jung, whether the attitude be one of agreement or disagreement.'

Theoretical Models and Personality Theory. Edited by D. KRECH and G. S. KLEIN. Durham (N. Carolina): Duke University Press; London: Cambridge University Press. 1952. Pp. 142. 19*s*.

Reviewing a symposium is often a puzzling task. One of its difficulties is that the individual papers must be assessed not only as separate items, but also as contributions to a whole—and a whole, moreover, whose aim may not be well-defined. In this volume a general discussion of theoretical models by von Bertalanffy is followed by papers in which Hebb, Rapaport, Miller, Eysenck, Halstead and Angyal describe and defend the type of psychological model which they prefer—the neurological, psycho-analytic, Hullian, factorial, organismic. What they have to say is in general well said, but in its broader aspects hardly novel; the positions are familiar, and one is led to wonder why it was thought worth while to bring them together under this title.

Some part of the answer is perhaps to be found in the opening paper by the editors, who deny that there is any distinct body of theory which can be called 'personality theory' and set slongside other subsystems such as 'learning theory' or 'perception'. They very correctly point out that the existence of these subsystems is itself a barrier to theoretical advance in psychology; and that personologists who, by adopting such slogans as 'personality and perception', perpetuate these divisions, delay the construction of the unifying framework of behaviour theory within which they should properly seek the answer to their own questions. Klein and Krech then attempt to state the requirements of such a behaviour theory.

The correctness of their requirements, and the extent to which any of the proposed models can meet them, are questions which might profitably have been taken up by the later contributors; as it is, they have simply been passed to the reader. Can the title of a symposium not be made to function as something more than a projective device?

N.E.L.

Affective Disorders. Psychoanalytic Contributions to their Study. Edited by PHYLLIS GREENACRE. New York International Universities Press Inc. 1953. Pp. 212. $3.00.

As Dr Greenacre tells us in a short Foreword, this book, which contains contributions by five other psychoanalysts, represents part of certain 'panel discussions' held at meetings of the American Psycho-analytic Association in 1950 and 1951. She regards it as 'a heartening indication of the growth of analytic vision that it again concerns itself with psychoses and borderline states, although these conditions are not generally amenable to formal psychoanalytic technique, but require special flexibilities, modifications and empathy'. The 'concern' may indeed be praiseworthy, but it seems doubtful whether the results so far obtained, to judge by the contents of this book, will be generally considered 'heartening'. The present discussions do not touch, at least directly, on therapy, but are concerned with theoretical formulations—in terms of 'metapsychology'—regarding the processes operative in depressive or manic states. The reader is left with the impression that little definite progress along these lines has really been achieved in recent years. The psychoanalytically orientated psychiatrist will probably be interested to see how far the very tentative formulations put forward by the discussants can throw light upon his own experience, but for most others the book may well appear both difficult and depressing. To them the writers may seem to be only too often engaged in rather unsuccessful manipulations of the concepts 'ego', 'id', 'super-ego', 'orality', 'pleasure principle', 'projection', 'introjection', etc., in the hope that they can be placed into such a relation to each other that they will explain what happens when a patient succumbs to melancholia, or mania, when he changes from one state to the other, or when he recovers. They appear to be grappling, as it were, with a very refractory jigsaw puzzle.

The difficulty is increased by the fact that the use of these terms is not always quite consistent or in accordance with tradition; as though indeed the pieces of the jigsaw were liable to change their shapes when handled. To give just two examples: On pp. 25–26 Dr Edward Bibring tells us that in his view depression is 'not determined by a conflict between the ego and the id or the super-ego... but stems primarily from a tension within the ego itself... since it feels unable to live up to its aspirations (ego ideal, superego) while they are strongly maintained'; here the words within brackets seem in contradiction to what goes before. On p. 146 Dr M. Katan says that 'the superego, as the representative of reality, is the most important pillar of the reality principle', and adds no explanation or qualification, in spite of the generally recognized existence of powerful archaic and unrealistic aspects of the super-ego, and in spite of the fact that some thirty years ago Freud in *The Ego and the Id* (p. 34) expressly repudiated his earlier view that the super-ego included the 'testing of reality' among its functions.

It is clear both from Dr Bibring's introductory paper and from those that follow that the writers no longer consider that oral fixation, incorporation of a lost love object, turning of aggression against the self, and ego-super-ego conflict, play quite such an important role in the production of melancholia as they seemed to do before, though there appears to be little general agreement as to what further explanations are required. Dr Bibring regards the ego's sense of helplessness as the essential factor. Dr Edith Jacobson is perhaps nearer the 'classic' view when she stresses what she calls the 'symbiotic' type of relationship—involving dependence on supplies of (oral) love—as predisposing to depression, though she also thinks that at the onset of psychosis proper there is a defusion of the libidinous and destructive forces, leading to a 'fatal libidinous impoverishment and an accumulation of sheer aggression'. She would seem to be thinking here in terms of Freud's Eros and Thanatos. Dr Elizabeth Rosenberg Zetzel, however, while accepting the importance of Melanie Klein's 'depressive position' in infancy, does not consider that this necessarily involves the operation of a 'death instinct'. Dr Gearge Gero, in a study of anorexia, is inclined to find a cycloid process in which the super-ego first curbs the longing for love, this being followed by an unavailing appeal by the 'depleted ego' for love from the super-ego itself; while Dr Katan, in considering mania, holds that the predominant working of the pleasure principle has a therapeutic value in the patient's effort towards recovery, by healing the 'narcissistic wound' and 'causing the surplus of id agression to disappear', though he is quite frankly vague about the dynamics of the processes involved.

The general trend of the discussion is perhaps best indicated by Dr Gero when he says that at the present time 'We are much less apt to reach easy generalizations but much more able to develop concepts and correlations which are descriptive of the unending varieties of the phenomena presented to us by our patients'. We must however, it would seem, resign ourselves to the fact that these concepts and correlations have as yet hardly advanced beyond the stage of trial and error groping.

J.C.F.

Psychoanalytic Studies of the Personality. By W. Ronald D. Fairbairn. London: Tavistock Publications. 1952. Pp. xi + 312. 25*s*.

The author presents a highly individual contribution to psychoanalytic theory under a misleadingly general title. He offers an object-relations theory of personality, ousting the Freudian libido from its fundamental place, and assuming, like Klein, a fundamental schizoid split in all personalities. Freud's libido as pleasure seeking is supplanted by Fairbairn's libido as object-seeking, and the concept of endopsychic structure which derives from this object-relations theory is conventionally tripartite. A central ego and two subsidiary egos (libidinal ego and internal saboteur) correspond superficially to ego, id and super-ego, but the ultimate difference lies in the fact that Fairbairn bases endopsychic structure upon repression of internalized objects. The important practical and theoretical inference is that the ego is inherent and that the subsidiary egos are derivatives from this original unitary ego, as opposed to the Freudian position of the development of the ego from the undifferentiated id. To Fairbairn, id, or libidinal ego, is a split-off portion of the original dynamic ego. Abraham's theory of libidinal development is replaced by a psychopathology of the psychoses and psychoneuroses using the processes of accepting or rejecting, internalizing or externalizing objects, and based on a scheme of development from infantile dependence through the transitional stage of quasi-dependence to mature dependence.

The theory stems from the author's sensitive clinical experience and insights. Thus his work, while mainly relevant to psychoanalysis, has something for all interested in the dynamics of personality. To the general psychologist, the formulation leading to the concept of a unitary, inherent or original ego is important, but the less specialized reader would be well-advised to provide himself with a background of neo-Freudian development before tackling this book.

Dr Fairbairn does himself an injustice in the presentation of his views in this book. He collects his articles and essays over the last twenty-five years from various sources and connects them by footnotes and addenda. This arrangement, however, permits too many minor contradictions, alternations and repetitions to distract the reader from the main theme of the book.

R. M. MOWBRAY

Aggression and its Interpretation. By LYDIA JACKSON. London: Methuen. 1954. Pp. 237. 21*s.*

The title of this book promises a more comprehensive treatment of the subject than the author in fact gives us. The main part consists in an account of two experiments and some therapeutic work with children in the phases of latency and early adolescence. These studies are prefaced by a sketchy discussion of theories of aggression. The vexed question to which the author addresses herself is that of the origin and transformation of aggression, her thesis being that 'aggression is a manifestation of self-assertion...grounded in the survival urge'. One may doubt whether any empirical studies can either confirm or refute such a broad quasi-biological hypothesis and, indeed, the link between the author's theory and her experimental and clinical observations is not a close one.

In the main study reported, projective tests of family attitudes were administered to normal, neurotic and delinquent children, the normal subjects giving a much more positive picture of family relations than either of the disturbed groups. The finding that the delinquents admit less hostility than neurotics to parents and sibling-figures, idealizing the absent family, is difficult to interpret since more of the delinquents than neurotics seem to have been living in institutions at the time.

The clinical studies of behaviour under the influence of a cathartic type of play therapy provide many interesting examples of overt aggression giving way to more constructive forms of relation, but the therapist's role is so vaguely described that it is hard to understand the dynamics of these changes.

In style the book is rather conversational and would have profited by a good deal of 'cutting'.

CECILY DE MONCHAUX

Aggression, Hostility and Anxiety in Children. By LAURETTA BENDER. Springfield, Illinois: Charles C. Thomas. 1953. Pp. viii+184, 23 illustrations. 40*s.*

This expensive volume is beautifully produced and leads the reader to expect much. In my opinion the expectation is not fulfilled; nevertheless a great deal of good clinical work goes to the making of the book and anyone studying the problems brought forward will not be able to miss reading it. There are seven separate studies written by Paul Schilder and Lauretta Bender and their associates at the Bellevue Hospital. These deal respectively with Aggression, Death, Suicide, Homicide and Firesetting, as these appear in child psychiatry. The genesis of hostility is discussed. There is a final chapter on Anxiety.

In the studies on aggression and death, projection tests are used. In the latter case there is a series of pictures and in the reviewer's opinion one of these is quite unsuitable for a test to be used with children. It shows a number of men and women hanging. One is used to every kind of horror that appears in the children's own drawings of their dreams and conscious fantasies, but it is quite another matter to present children with reality horrors which many mature adults must find disturbing. It is a sign of the times, one supposes, that the official description of this picture is 'Russian spies hanging from gallows'. In this country both the picture and the description would be wholly out of place. I hope it will be understood that this criticism does not imply any scotomization of the capacity of the child both normal and abnormal to conceive of the horrible.

In each section rich and interesting surface descriptions are given and these can have value to those who are not able themselves to be in touch with clinical material. Anyone clinically involved would of course derive more from the careful study of a few personal cases than from the perusal of a large collection of cases described by other observers. What is needed from this book is new light on the origin of aggression. Readers will seek for this in the last two chapters where American authors are well quoted and also Melanie Klein. In the reviewer's opinion there is no possibility of a valuable new idea in this book because of what seems like determination on the part of the authors to avoid a developmental approach. Quotations from Freud and Melanie Klein particularly show the way in which statements are taken without a full appreciation of their meaning in relation to development.

A comment on Melanie Klein's theory on page 145 that anxiety springs directly from aggression which the authors find 'conclusively disproved' seems to show no knowledge of Klein's complex statement of the depressive position in normal emotional development. In place of the general psychoanalytic statement of the developing individual, one gets a general theory that aggression in childhood is a 'symptom complex resulting from deprivations which are caused by developmental discrepancies in the total personality structure, such that the constructive pattern drives for action in the child find inadequate means of satisfaction....' Here the word development occurs but the idea does not seem to be followed through.

The best part of the book is an unexpected section at the end on schizophrenia in childhood. Here however the operative factor seems to be the birth experience. If the authors become able to make a statement of *what the infant is like* when affected by birth experiences they will then be able to re-state their theories in a more convincing way.

D. W. WINNICOTT

Child Training and Personality: by JOHN W. M. WHITING and IRVIN L. CHILD. New Haven: Yale University Press; London: Geoffrey Cumberlege: Oxford University Press. 1953. Pp. viii + 353. 32*s*. 6*d*.

Readers of Dr Whiting's earlier work may well be puzzled and perhaps disappointed by the present book written in collaboration with Dr Child. It is an extremely ambitious cross-cultural study which examines child-training practices in seventy-five societies, relating the resultant personality types with various aspects of social organization. The method of approach is psychoanalytic, which has much to recommend it, provided those who employ it recognize that other branches of social science have something to add to the understanding of society which can be gained by psychological techniques. But the effective use of psychoanalytic concepts surely necessitates, first and foremost, a large amount of accurate information about the individuals and groups studied, and the ethnographic material about these seventy-five societies varies enormously in quantity and quality. There is, for instance, a great deal about the Navaho and Zuni, while the material on the Abipone was published in 1822 when modern ethnographic methods were non-existent, and the material on the Lapps consists of Turi's *Book of Lappland*, which, while fascinating reading, is hardly a book of anthropological significance, and Scheffer's *The History of Lapland* published in 1704. Moreover, even where the material is modern and reliable, the difference in approach and interest of various anthropologists makes it a risky procedure to compare their findings on topics which they may have regarded variously. We would, for instance, discover little about the metabolic processes of the human body by studying reports by a psychoanalyst on an Australian aboriginee, and by an Ear, Nose and Throat Specialist on an Eskimo.

But there is a further difficulty about the present study. The authors' conclusions are very largely reached on the basis of assessments by three judges (a logician and two psychologists) who have read the ethnographic material. These persons have been asked to 'evaluate the general effect on the children in a given society of all types of punishment and reward used in weaning, as he can judge it from the entire context of description of child training procedures in that society'.

Needless to say, the method is not quite so fantastic as this isolated quotation would suggest. A certain degree of statistical sophistication has gone into the correlation and evaluation of these judgements, while the judgements themselves are broken down into a large number of more specific questions. But no amount of ingenuity can quite conceal the unreality of the whole procedure. We know little enough about what happens in a single society for which the most complete first-hand evidence exists, and it is surely both premature and misleading to start system building at this stage of our ignorance.

However, notwithstanding the use of methods which cannot be allowed to go unquestioned, there is much which is stimulating and interesting in this book. The authors are men of sensitivity and wide experience who, however slender their grounds, have penetrating and challenging things to say on the general questions of culture-personality studies. They conclude their book by calling for more and fuller studies of society, and if they were not prepared to wait for these, they have at least illustrated our need for them.

ADAM CURLE

Phantasy in Childhood. By AUDREY DAVIDSON and JUDITH FAY. London: Routledge and Kegan Paul Ltd. 1952. Pp. viii + 188. Price 18*s*. net.

As the cover states, this work is for a psychoanalytically oriented public. The authors follow the Kleinian school. They have had a personal analysis, and, between them, experience in a children's club, child guidance clinic, hospital, nursery school, residential nursery for infants. And they worked with homeless children and with individual neurotic and psychotic children.

They show the kind of unconscious phantasies which operate at the various Freudian levels of development, and, during middle childhood, how these phantasies become modified by the experience of the actual world in which the child lives. In the last chapter, 'Living through of phantasies', an account is given of the working out of some phantasies in the life of a single child up to about three years of age.

The style is easy and flowing, and general statements are freely illustrated from particular children who are introduced by name straight into the text when it is appropriate. This makes for continuity of reading, but if anyone suddenly wonders, for example, which child is 'Charlie', then he will have some difficulty in finding 'Charlie' in the previous pages since he is so embedded in the text. Unfortunately there is no Index.

If the reader likes this way of presentation, which is a clever dovetailing of the theoretical with the concrete, he should find the reading a kind of living experience. There is more in the book to be pondered over than there appears to be, just because the style is so easy and the language unpretentious. It gives many vivid glimpses into the phantasy life of the child, and shows among other things how helpful and at times important it might be for a child that an adult should respond to this inner life to help him to a new understanding of himself, which in turn would enable him to cope better with life. The book is not dogmatic. Sometimes interpretations are merely suggested, while other interpretations must be based on much material not actually set down in it. Part of the value of the book lies in its suggestiveness.

A.S.P.

Understanding Children's Play. By RUTH E. HARTLEY, LAWRENCE K. FRANK and ROBERT M. GOLDENSON. London: Routledge and Kegan Paul. 1952. Pp. xvi + 372. 21*s.*

This book describes a project carried out on children of two to six years of age, mostly in nursery schools. Different play activities were investigated to find out how far one could understand children from their play and in what ways they could make use of different play situations to express and to discover themselves. The conditions were purposely left uncontrolled but adapted to fit into each school programme and therefore varied from school to school and teacher to teacher in the freedom allowed to the children.

Each material or situation is discussed in a separate chapter: dramatic play, blocks, water, clay, drawing and painting, music and movement. A description is given from records taken at the time of what the children do with each means of expression and what it seems to do for them. Some case histories are given to show the relation between the different kinds of play and personal problems, and some estimate is made of the therapeutic value of play for children with difficulties. It is pointed out that more varied opportunities of this kind could be given in the ordinary school system since this report shows that with courage and thought even water-play can be arranged. To those educationists who may think life should consist of hard work it may be remarked here that children often work hard at play but this task is more often self-chosen and self-imposed.

Too much weight perhaps has hitherto been given to intellectual, moral and physical disciplines throughout our school education, and too little has been done to satisfy the emotional needs of the child for his balanced development; too much emphasis on learning things and too little upon experiencing them.

Much attention has been drawn to the use of play as a method of teaching young children and as a diagnostic and therapeutic method in clinical work. We must thank the authors of this book for carefully examining the process of play itself and showing that it could be a valuable part of the school life of the ordinary child. It serves, of course, a similar function to play outside school but differs in effect, chiefly because it is performed in what one might call a sheltered situation. There are an Appendix of Suggestions, Notes to Chapters, a Bibliography and an Index.

A.S.P.

Practical Child Psychotherapy. By CURT BOENHEIM. London: Staples Press. 1953. 2nd ed. Pp. 184. 15*s.*

This book was first published in Great Britain in 1938. As its subtitle states, it is primarily a guide for the general practitioner. It discusses under the headings of certain physical symptoms or disorders in children such as loss of appetite, enuresis, asthma, to what degree they may be brought on by emotional stresses in the family situation, to what extent they should be treated, in addition to any suitable physical remedies, by psychological methods. These may vary from the simplest kind which any good physician would probably use, to cases which must be sent to a professional child psychotherapist. The author has drawn on his sixteen years' experience as a child psychotherapist in Berlin. The majority of references in bibliography are to German texts, but there are also references to a number of English ones. A.S.P.

Early Phases of Personality Development: a Non-formative Study of Infant Behavior. By SYBILLE ESCALONA, MARY LEITCH and others.

Monographs of the Society for Research in Child Development, Inc. Volume XVII, Serial no. 54, no. 1. 1952. U.S.A.: Child Development Publications. 1953. Pp. vi+72. No price given.

This monograph describes research aims and methods applied by the authors, their collaborators and assistants in the study of 128 infants below seven months of age. No definite findings or conclusions are so far presented, but certain further researches to be undertaken by various members of the team are outlined. These studies are all to be based on the same quite extensive data.

The work was undertaken by a total team of seven psychiatrists and psychologists, all engaged in clinical work at the Menninger Clinic. A staff of five secretaries kept the records. The interweaving of clinical and research activities, based on empirical studies of infancy, should be productive of useful results.

The primary interest of the team was in personality development and the approach was in the first place psychoanalytic. The study is described as non-normative; but, far from being merely descriptive as might be expected, many techniques including quite extensive measurements have been incorporated into the study. The interests apparently range from physical and anthropological measurements, medical investigations, psychological test results, environmental studies, psychiatric syndromes, modes of normal infant behaviour functioning and their variability, emotional factors, to mother-child relationships, etc.

Methods used include interviews with mother and child in a specially and quite elaborately prepared centre, detailed observations of infant behaviour in a variety of situations, home visits, cinematograph records, tests of intelligence, physical examinations. The psychological tests used were the Cattell Infant Intelligence Test, and the Gesell Developmental Schedules. Special observational procedures were also undertaken: tissue resistance, leg resistance, strength of grip, etc. A stabilimeter was used to measure (by means of kymographic tracings and counter readings), the baby's movements in greater detail. Activity ratings, for example before and after feeding, were assessed. Two films were made of each baby to show feeding behaviour and other spontaneous behaviour. Running records and notes were also kept by additional observers.

The result of all this work, which has already occupied a number of years, is a very extensive mass of data that the authors expect will take several further years to organize and summarize. A form consisting of no less than 40 pages of headings has been prepared for the summarizing of the material in each case. The studies which it is proposed shall result from the data include: (1) exploration of infantile functioning in three body systems—the sensorium, the skin and the gastro-intestinal; (2) a survey to study whether behaviour syndromes can be delineated; (3) maternal attitudes and determinants of specific attitudes in infants; (4) sleep patterns; (5) maternal behaviour in feeding infants; (6) psychological test findings and 'adjustment status' in infancy; (7) a study of tact; (8) traditions about infants and infant management; (9) hand and arm movements and postures in infants; (10) the emergence of voluntary, purposeful action.

This piece of work as described should prove of interest to a wide audience, as it attempts to cover or touch on an enormous number of facts and topics of medical, psychological and psychoanalytic interest. This is perhaps also its weakness, that it has attempted to do so much that it may prove difficult to crystallize out information that is new or significant. Starting as a 'non-normative study', it appears that norms of many kinds have yet been introduced as the work proceeded, and the resulting data should provide a valuable core from which useful and exact information could be extracted. It will be of great interest to note what use is made of this material, both normative and descriptive, in the contemplated researches.

RUTH GRIFFITHS

The Abilities of Babies. By RUTH GRIFFITHS. London: University of London Press. 1954. Pp. x+229. 20*s.*

In this book the author has tackled the difficult task of evolving a scale for measuring the mental development of babies up to two years of age. It is written in three parts. The first, based on careful observation of babies, and full consideration of existing tests, points to the possibility of a new scale with a new technique for assessing the abilities of this infant group. The second part details the seeking, searching and researching necessary in order to allow of the building and presentation of such a scale. The third demonstrates the method of administration of the new scale which has been founded on aspects of development that appear significant for mental growth. In its published form the scale consists of five subscales lettered *A* to *E* and all very nearly equal in difficulty. Each subscale covers one vital aspect of infant development namely: *A*, locomotor development; *B*, personal-social adjustment; *C*, hearing and speech; *D*, eye and hand co-ordination; and *E*, performance.

From a summation of the scores from these five subscales Dr Griffiths gives a means of obtaining a G.Q.—General Intelligence Quotient. Apart from this G.Q. it is also possible to calculate *QA*, *QB*, *QC*, *QD* and *QE*, each representing respectively a quotient indicative of the child's ability in each of the five subscales. Since these five scales are very nearly equal in difficulty it is possible to show the total mental level attained by a subject at the time of testing. This has been done by the profile method, which from the diagnostic point of view should prove most valuable in that it depicts general or specific disabilities or the presence of exceptional ability in one or more of the basic developmental processes.

The inclusion in the book of developmental profiles of particular cases, and of excellent photographs of the responses of babies to different situations, adds to the clarity of the exposition. There is an excellent bibliography. Methods of scoring the tests and assessing results are assembled with admirable economy and clarity in Appendix I.

The Abilities of Babies will be of interest and of value to all concerned with the health and well-being of children, and particularly to the psychiatrist and the psychologist working in the field of child psychology. It offers a new approach to assessment and to diagnosis. It suggests many new avenues for further research. This scholarly work commands respect for its clear exposition of each step in procedure, its integrity of observation and report, and its wealth of detail which illumines the text without tedium. Young research workers will find research procedures fully mapped out.

Dr Griffiths intends to 'continue the work on the new scale beyond the second year and this will then make differential diagnosis as outlined in this book available for testing children over two years of age'. We look forward to the completion of this further research. Such a scale carried to years five or seven would prove invaluable, particularly in the field of the handicapped child. MARGARET MARTIN

Childhood Problems and the Teacher. By C. BÜHLER, F. SMITTER and S. RICHARDSON. Routledge and Kegan Paul. 1953. Pp. xii+262. 25*s*.

Educating the Subnormal Child. By F. LLOYD. London: Methuen. 1953. Pp. vii+148. 11*s*. 6*d*.

The Education of Exceptional Children. A. O. HECK. 2nd ed. New York, Toronto and London: McGraw-Hill Publications Inc. 1953. Pp. xiv+513. $6.00.

These three books on child psychology are concerned with children who are deviants from the normal.

The first of these, as its title suggests, is written specifically for the teacher and attempts to show how everyday problems in the school may be solved by the teacher. The book is the work of three authors who are experts in the field, a fourth writer contributing a chapter on remedial teaching. Children's drawings illustrate the text and their interpretations give excellent insight into the difficulties and problems of their respective creators.

The majority of the disturbances which arise in school or in class need to be treated only on the surface, but others may require deeper psychological study. Symptoms of such disturbances are discussed and it is suggested that the teacher who has sufficient psychological knowledge and has made a study of personality problems may be able to deal with some of these. Others require the more expert advice of psychologist and psychiatrist.

Part I deals with the psychological approach to children's problems. Certain sections, which consider basic concepts, appear to be too summarized for easy reading, and presuppose some knowledge of psychology. Part II, concerned with what the teacher can do, forms the main part of the book and contains discussions of case histories. Part III is an assessment of the psychologist's collaboration with the school and is concerned with diagnosis and psychotherapy. It also contains a chapter which serves as an introduction to projective techniques.

The book draws attention to problems met with at school and should be of interest to intelligent parents and teachers. It presupposes, however, that teachers have a fairly extensive knowledge of clinical work, have access to the child's home, and that parents always co-operate.

The author of the book on educating the subnormal child is in charge of a school for subnormal children, and the content of the book is an attempt to convey to the reader her attitude to the children under her care. Individual understanding is the main theme and examples are given of children who show all-round improvement in an environment in which they feel secure.

This is not so much a book for the professional psychologist as for the special school teacher, whom it will stimulate to further endeavour.

The volume from the hands of Prof. Heck has brought up to date an account of the education of exceptional children first published in 1940. One chapter has been wholly rewritten and sections and paragraphs in others have been revised. Although the setting is in the U.S.A. the contents of the book should be of value to all those concerned with exceptional children. The contents are divided into four parts, the first three considering in turn the education of the socially handicapped, the physically handicapped and the mentally exceptional. The term exceptional not only applies to all those who are handicapped, but also to those who are gifted, either generally or in one specific direction. The fourth part discusses problems of administration. Photographs are included which illustrate new equipment and new procedures, and for those interested in any type of children who require special education the book should prove stimulating. It is sufficiently non-technical to appeal to the non-specialist, although it was written for use in college and university courses.

MARY COLLINS

Family Environment. By JOHN D. NISBET. London: The Eugenics Society and Cassell and Co. 1953. Pp. 51. 3*s*. 6*d*.

A negative correlation between family size and intelligence has been reported by many writers, even after environmental factors such as occupation have been eliminated. In this study the author gives a carefully reasoned account of the many problems involved in assessing the relationship; he gives an extensive survey of the literature on the subject and also discusses the verbal factor in mental development. He then puts forward the hypothesis that a large family is in itself deleterious to the mental development of a child, particularly in so far as it affects verbal development; in support of this hypothesis he presents the results of a survey carried out on Aberdeen school-children. These results are certainly of interest, but his conclusions from them are not altogether convincing owing to a major fault in an otherwise carefully planned survey. The correlation between family size and certain test results which he has calculated are obtained from all Aberdeen schoolchildren in certain age groups and no account has been taken of varying social class or occupation of parents, which is known to be related to family size. Unfortunately these factors could very well account for his results, so that his conclusions are almost entirely vitiated.

M. A. CREASY

Social Implications of the 1947 Scottish Mental Survey. London: University of London Press for Scottish Council for Research in Education. Pp. xxiii + 356. 10*s*. 6*d*.

In the second volume on the results of the 1947 Mental Survey of Scottish 11-year-olds Mr James Maxwell makes a careful and thorough analysis of the sociological data collected about a random sample of 7380 children from the total group of 75,211, namely those born on the first three days of each month of 1936. This analysis shows that there is a steady decrease in average intelligence test score, height, weight, frequency of migration from district to district and age of mother as one ranges from professional and large employers through salaried employees, small employers, non-manual wage-earners, skilled and semi-skilled workers to unskilled manual workers; and at the same time a steady increase in family size and overcrowding. *Within each occupational class*, however, average test score shows a definite downward trend as family size increases. No marked differences were found between the average test scores of children living in 'Cities', 'Large Towns', 'Small Towns' and 'Other Areas'.

Mr Maxwell frankly admits that the Survey 'raises more problems than it solves'. It offers, for instance, no further light on the problem of the relative contributions of heredity and environment or of the different factors within the environment to the better (on the average) performance of the 'upper class' child from a good home on an intelligence test, but it is 'fairly definite' in its conclusion that 'intelligence, as measured by the test, does vary with environmental circumstances to a considerable degree'.

Included in the volume is an analysis of the test scores of 947 twins from the *total* survey group. This confirms the observation of other investigators that twins are on average four or five points of I.Q. below non-twins, but throws no further light on the reasons for this difference.

ALEXANDER LAING

Education in England. By W. P. ALEXANDER. London: Newnes Educational Publishing Co. Ltd. 1954. Pp. xi + 147. 12*s*. 6*d*.

The subtitle of this book is 'The National System—How it works'. This gives the aim which Dr Alexander had before him. In the foreword he comments on the difficulties encountered by visitors from abroad in trying to grasp how the English educational system is organized. The book attempts to meet these difficulties, and no doubt foreign visitors—and our own students—will be grateful for this attempt to fill the gap in our library shelves. The absence of any previous volume of this kind is perhaps an indication of the difficulty of the task, and it would be easy for educationists to criticize the book for what it omits—with such a complex subject this is inevitable in a work of this size.

The reviewer was surprised at the absence of any diagrams and 'flow-charts' which have the advantage of presenting complexity succinctly, and which (particularly for the foreigner) are sometimes less subject to misinterpretation than the many paragraphs of print necessary to outline the same ideas verbally. Dr Alexander is a distinguished administrator, and one feels that the book presents the machinery of the system without sufficiently indicating the parts of the machine that work smoothly and those that creak and groan a little. The foreigner may well feel frustrated over some aspects—for example, special schools for the educationally subnormal are given a subsection, yet the criteria adopted for the ascertainment of the children are not referred to, nor is any indication given of the numbers involved. The Ministry pamphlet covering this point is mentioned in the text, but its full title is not given, nor is it included in the appendix of references. The absence of an index to the book is particularly to be deplored, even though the table of contents is given in detail.

These criticisms, however, do not affect the judgement that the book as a whole is one which fills a real need of the student and the visitor from overseas.

STEPHEN WISEMAN

Psychology Applied to Human Affairs. By J. STANLEY GRAY. New York, Toronto and London: McGraw-Hill Publications Inc. Second edition. 1954. Pp. viii + 646. 42*s*. 6*d*.

This second edition of the 1946 publication again contains seven chapters by the author-editor, together with chapters from ten other contributors. It is a text-book of applied psychology, designed to follow a short introductory course of general psychology. The following fields are covered—college life, child development, educational psychology, vocational guidance, personality adjustments, speech correction, mental illness, clinical psychology, effects of diet and drugs, crime and crime detection, psychology in industry (four chapters, including a new one on water determination), marketing and advertising, military psychology, music and art, public opinion and propaganda.

The general impression of this slightly more up-to-date edition is that its claim to be a bird's-eye-view only, is fully justified. The author points out that it needs a teacher 'to interpret facts and answer questions'. Certainly much of the material is over-concise and could be misleading. Apart from some useful hints to students of the mental hygiene sort, a course in applied psychology would still be more usefully based on existing texts of a less comprehensive kind.

If one were to judge from this book, almost no applications of psychology to human affairs have been made in Britain. Only a very few articles mostly published in the twenties, are mentioned.

It is a pity that a costly book of this length should be so unrewarding.

FLORENCE MACNEILL

Transactions of the Conference on Morale—And the Prevention and Control of Panic. New York: 1954. Pp. vi + 75. No price given.

This impressively titled report was no less impressively sponsored—by the New York Academy of Medicine and the Josiah Macy Jr. Foundation. The 75 pages of typescript contents hardly bear such a weight. There are seven papers: J. A. M. Meerloo, 'People's Reaction to Danger'; G. M. Gilbert, 'Social Causes contributing to Panic'; H. Brucker, 'The Press in relation to Morale and Panic'; R. Saudek, 'Radio in relation to Morale and Panic'; G. S. Stevenson, 'The Voluntary and official agencies Role in the Promotion of Morale and the Prevention of Panic'; J. S. Tyhurst, 'Research on Reaction to Catastrophe'; and L. Alexander, 'Morale and Leadership'. In a book planned for publication and serious criticism this combination of authors and titles would have been promising.

Unfortunately we are given only the verbatim report of short speeches with these titles, given to what must have been a very mixed audience. At least, the speakers, with few exceptions, took no risks and preferred the colloquialism to the precise statement, the anecdote to the reference to research, and the extended cliché to the novel insight. The information we have gained from MacCurdy's 'Structure of Morale', the wartime collection of papers on 'Civilian Morale' edited by Goodwin Watson, and the vast series of attitude surveys carried out by the Information and Education Department of the U.S. War Department, is neither usefully summarized nor valuably extended. One is indeed struck by the combination of statements emphasizing the importance of research into problems of morale and panic with an apparent ignorance of even the best known work carried out in this area. At least one of the speakers, however, was concerned solely with the problem of psychiatric 'first aid' in a situation following a catastrophe, and Dr Meerloo in an admirable paper gives a series of hints for the mass treatment of acute anxiety states.

In general, one's feelings about such publications are mixed. Conferences attract publicity, and a glance inside the bright red covers of this report may have directed the attention of some foundation executive to the research needs of the social sciences. But the gulf between the spoken and the written word is too great, at least on this occasion, for one's reading to be very enjoyable or fruitful. Perhaps the predominating impression is best conveyed by Dr Iago Goldsten in his concluding remarks, when he says 'I should here give a résumé, but I am afraid I shall fail you on that score. The word résumé, as you know, means to resume. To resume means to begin all over again and on this occasion to begin all over again would indeed be terrible.'

J. F. MORRIS

The Nature of Choice in Casework Process. By ANITA J. FAATZ. North Carolina University Press; London: Geoffrey Cumberlege. 1953. Pp. xii + 141. Price 24*s.*

This book discusses the freedom of action and behaviour of the client of the social worker from the point of view represented by the author. This point of view found its expression in the will psychology of Otto Rank, who has had such great influence in various directions in the U.S.A. but is less well known in Britain.

The discussion is preceded by a historical account showing how the concept of the freedom of the individual first arose in connexion with social work and has been stressed as important from time to time from 1880 onwards, and has changed in meaning.

The main part of the book shows what is this freedom (nature of choice) a client should enjoy. It is assumed that anyone who comes to a social agency for help—no matter how simple—has projected his own inner difficulties upon the outer world. Those difficulties with which the individual is striving are the concern of the social worker, within the framework of the practical problem brought to him.

In order to explain how the social worker can help his client with these difficulties and yet leave him free each time to make his own choice, it is necessary to explain the whole 'functional helping process', which involves explaining the caseworker and his development in experience and depth of understanding and the relationship between client and worker. This is done step by step until we arrive at something very like the psychoanalytic process very ably described, partly in Rankian terms and partly in those suitable to the context of social work and the author's mode of thought. The determinism of Freud is repudiated and the freedom and responsibility of the individual are postulated. At each stage he is free to choose his actions and he is responsible for his choice. The social worker 'stands by', nautically speaking. In this context the desirable behaviour is to take (to will) the choice which leads towards the growth or maturity of the self. Nevertheless, the crisis in the 'helping process' comes when the client accepts the inevitable and sets out to achieve a new relationship to it. Although this is described in a rational way the process itself is not primarily a rational or intellectual one. The 'will' is described as 'the expression of the creative self'.

The account is very abstract, and the American technical or semi-technical language is at times slightly repellent to the English ear, probably because we are inured to our own jargon. The author has had much experience and has thought deeply before writing this book. It sets out a philosophy and a goal for social workers on the therapeutic side of their profession, and would also be of value to anyone engaged in dealing with other people in an advisory or helpful capacity, and who wants to understand what may lie below the superficial level of the immediate situation.

A.S.P.

STANLEY J. F. PHILPOTT, 1888–1952

OBITUARY NOTICE

STANLEY J. F. PHILPOTT

1888–1952

In Dr Philpott the British Psychological Society found one of its most devoted and ablest servants during its formative period. From 1929 until his death he gave himself wholeheartedly to its service in a variety of offices. For the greater part of this time he was Treasurer, and did more than any other to establish the finances of the Society. He was an extremely good man of business, one who could keep details constantly in mind, and one who was very firm on routine without lacking the constructive imagination to deal with new problems. Only those who had the privilege of then working with him can understand how much the Society owed to him during the lengthy preparations for incorporation. We faced the complicated task of drawing up a constitution which would satisfy the Society's members and the Board of Trade, and he brought to it a finely logical mind and considerable powers of negotiation and argument. It was he who in the end succeeded in convincing the Board of Trade that we and no others had a right to the title we had already used for a quarter of a century. In 1948 he was elected President. He served psychology in other spheres. He served the full range of offices in Section J of the British Association, and was a member of the Council of the Association at the time of his death. Further, he was associated with *The British Journal of Educational Psychology* from the beginning.

Philpott was trained as a teacher at Exeter, and he retained some touch with schools until the end. He lectured to various training colleges, and took part in pupil selection, employing here Intelligence Tests of his own construction and standardization. He was destined to teach. 'Once a teacher, always a teacher', he said, and so not even the prospect of more time for his researches could lure him from teaching. He won the devotion and affection of his students, for, as one of them (Dr F. W. Warburton) says, he thought of them first and foremost as people. He was able to induce them to undertake with real enthusiasm some of the most tedious experimentation that ever psychologist engaged in. By their aid he developed his original work on the rhythmical fluctuations of mental work. There was an enormous volume of work carried out under his direction: its scientific value cannot be discussed in this article.

He devoted the most vital years of his life to the study of 'waves'. His monograph on 'Fluctuations in Human Output' (*British Journal of Psychology: Monograph Supplement*, no. 17, 1932) laid the foundations for what followed. By laborious experimentation and no less laborious theoretical analysis he claimed to have established a fundamental psychological law, that mental efficiency fluctuates in a complex system of waves whose lengths have to be stated in logarithmic time units. Looking back over the years that followed its publication, it must be said that the psychological world has avoided rather than challenged his conclusion. He, with the patient determination that characterized him, tried over and over again to drive his point home, but it was too novel and too abstract to be either criticized in an informed manner, or accepted. It may be that Philpott, brooding over his concepts for so many years, came to see them as simpler than they were, and so failed to make concessions to other minds: perhaps he prematurely made known some of the consequences he drew from the basic theory, and thereby started hares which diverted his readers from the real trail. Certainly even those kindly disposed to him and his ideas found it all very difficult (though it is reasonably simple to those who are content to take it as simple) and were tempted to give it up. Others just ignored it.

It is beyond question that Philpott was deeply hurt by this attitude, though never embittered. He continued persistently to seek for criticism (and he was never one to regard himself as inevitably in the right), and when he failed to obtain it, he would put in a few days of strenuous gardening, and then return to the attack. To his very great joy, and the joy came all but too late, the *British Journal of Psychology* devoted in 1952 nearly a full issue to a discussion of his views. Dr Lewis Richardson, an old friend whom he greatly admired, rose to the attack and, in his happiest hours, Philpott replied vigorously. But even here it would seem that Richardson's criticisms centred more on the secondary, and possibly vulnerable, deductions from the theory than upon the foundations. So his life-work remains unassimilated, awaiting perhaps destructive criticism of the experimental technique or of the statistical analysis: it cannot, surely, be always left in timid neglect. Philpott came to psychology from the physical sciences, and he had little sympathy with those who excuse the shortcomings of psychology by the complexity of its subject-matter. He knew that physicists were attacking greater and more abstract complexity with greater boldness, and he was not afraid to commit himself in like manner.

Philpott was a most lovable man. Life had never been easy for him, but he never lost courage, nor cheerfulness, however great the burden of work he had to carry. He judged others shrewdly, but seldom failed to find good qualities in them. He was a friend to be valued. I have spoken of him as psychologist and as one who worked so well for the Society, but to his friends he was above all else just a fine man, a simple-hearted gentleman. The outstanding memory which those friends retain of him is his complete integrity, in every sense which that word can bear.

A. W. WOLTERS

THE CONCEPT OF HEED

By U. T. PLACE

Department of Philosophy, University of Adelaide

I. *Introduction* (p. 243). II. *The concept of heed* (pp. 243–244). III. *The contemplative theory of heed* (p. 244). IV. *Ryle's objections to the contemplative theory* (p. 245). V. *The dispositional theory of mental concepts* (pp. 245–246). VI. *Ryle's application of the dispositional theory to heed concepts* (pp. 426–248). VII. *Ryle's account of the logic of heed concepts* (pp. 248–249). VIII. *The objection to Ryle's account of the logic of heed concepts* (pp. 249–250). IX. *The dispositional theory restated* (pp. 250–251). X. *The case for the traditional view* (pp. 251–254). XI. *Conclusions* (pp. 254–255). *Reference* (p. 255).

I. Introduction

Do the words and expressions which the subject uses when he makes his introspective report, refer to internal events going on inside him? If they do, it is difficult to see why we should not use the subject's statements in order to formulate and verify hypotheses about such processes. If they do not it is difficult to see what reason we should have for believing in the existence of the sort of events which are described in the text-books of introspective psychology. If, as is suggested in this paper, some of them do and some do not, it becomes extremely important for the psychologist to be able to discriminate between the two cases.

Now in so far as the language of the introspective report is the 'psychological' language of ordinary speech, this is the question which has recently been exercising the minds of the philosophers; and in at least one case (Ryle, 1949) the conclusion that has been reached is preponderantly negative. In his book *The Concept of Mind* Ryle has attempted to show that the traditional view which holds that mental states and processes are private internal occurrences within the individual is mistaken. He does not deny that some of the statements which we ordinarily make about people, refer to states and activities of the individual that are 'private' or 'covert' in the sense that only the individual himself can report their occurrence. He would maintain, however, that such statements constitute only a small minority of the statements we make about our own and other people's minds.

I shall argue in this paper that the number of mental concepts which do entail a reference to covert states and activities of the individual is much larger than Ryle is prepared to admit. In particular it will be contended that a reference of this kind is involved in our ordinary use of such expressions as 'being conscious of' or 'paying attention to something', 'observing', 'watching', 'looking', 'listening', 'seeing', 'hearing', 'smelling', 'tasting', 'feeling', 'noticing', 'perceiving' and 'recognizing'.

II. The concept of heed

The key notion in Ryle's account of what is traditionally referred to as our 'apprehension of the external world' is the concept of 'heed'. As he himself points out (p. 136), this notion of 'heed', or 'heeding' is closely related to the concept of 'consciousness' which is the basic concept in all the traditional theories of mind. While heed does not carry quite the same theoretical load as does the notion of consciousness in the traditional theories, it is employed by Ryle in his analysis of a wide range of mental concepts, and a large part of his case against the view that mental concepts entail a reference to covert states, processes and activities, would seem to depend on his ability to show that paying heed is not a covert activity.

Ryle defines the notion of 'heeding' or 'minding' as embracing such concepts as 'noticing, taking care, attending, applying one's mind, concentrating, putting one's heart into something, thinking what one is doing, alertness, intentness, studying and trying'. Concepts which 'entail, but are not entailed by, heeding' include 'enjoying, disliking, pondering, searching, testing, debating, planning, listening, relishing, calculating and scrutinizing' (p. 136), looking (p. 232), observing, watching, descrying (p. 207) and recognizing (p. 223). Remembering something, according to Ryle (pp. 91 and 137–9) involves having paid heed to it at the time, while being conscious of sensations in one's body or objects in one's environment is evidently synonymous with heeding or noticing them (pp. 157–8). It will be seen from this that Ryle's concept of heed corresponds more closely to the traditional concept of attention than to that of consciousness. On the traditional theories consciousness is the basic notion of which attention is a special active or conative form. For Ryle on the other hand, it is 'attending' or 'heeding' which is the basic concept, and no distinction is drawn between paying attention to something and being conscious of it.

III. THE CONTEMPLATIVE THEORY OF HEED

The traditional or, as Ryle calls it, the 'contemplative' theory of heed or attention and consciousness in the form in which I wish to defend it, may be stated as follows. The expression 'paying attention' refers to an internal activity of the individual presumably of a non-muscular variety whereby he exercises a measure of control over the vividness or acuteness of his consciousness of (*a*) the sensations to which he is susceptible at that moment, or (*b*) such features of the environment as are impinging on his receptors, without necessarily adjusting his receptor organs or their position in any way. In paying attention to something the individual is regulating the vividness of his consciousness of the object or sensation in question and hence the number of its features of which he is conscious. The expression 'being conscious of something' refers to a peculiar internal state of the individual which normally accompanies any reasonably intense stimulation of his receptor organs, the particular form assumed by the individual's state of consciousness at a given moment being determined by the pattern of physical energies impinging on his receptor organs at the time.

Being conscious of something is by definition a necessary condition of the individual's being able to give a first hand report on that something either at the time or later. It is not, however, a sufficient condition of the individual's ability to make such a first hand report, since it is possible for someone to be conscious of things which he cannot put into words, without his actual capacity to verbalize being in any way disturbed. Likewise, though here the relationship is probably contingent rather than necessary, the successful performance of any skilled activity depends to a greater or lesser extent on the individual paying attention to, i.e. maintaining a vivid consciousness of, relevant features of the situation and his own activity with respect to it; but the mere fact that someone is paying attention to what he is doing does not entail that the performance will be adapted to the demands of the task.

IV. RYLE'S OBJECTIONS TO THE CONTEMPLATIVE THEORY

Ryle's first objection to this type of theory (pp. 136–7) is that it leads to a *reductio ad absurdum* in those cases where we speak of watching carefully or attentively. He points out that it is always possible to ask of a spectator, whether he has been a careful or a careless one. In order to interpret this on the contemplative view, he suggests, we should have to postulate an additional process of watching his watching, which is present in the careful spectator and absent in the careless one. This interpretation leads to an infinite regress, since it would always be sensible to ask whether or not this watching of one's watching was done carefully or not. There is, however, no reason why the contemplative view should force us to adopt this particular interpretation. As Ryle himself points out (p. 136), minding can vary in degree. There is, therefore, no reason why we should not say that the difference between the careful and the careless spectator lies in the amount of heed that each pays to the scene before his eyes. The careless spectator is not one who fails to watch his watching, nor is he completely oblivious of what is going on, he merely pays insufficient heed to it. What distinguishes the careful spectator from the careless one are the detailed and accurate reports which he is able to furnish as a result of the richness and vividness of the impressions with which his more active heed-paying provides him.

Ryle's second objection to the traditional theory of attention is that it fails to account satisfactorily for those cases where we speak of applying our minds to some task, such as whistling or driving a car. In this case, he argues (p. 138), we are not doing two things, whistling and minding, driving the car and attending to our driving; we are performing a single activity in a certain way. He points out in support of this contention that we cannot stop driving the car and continue our heed-paying. This argument, however, is singularly unconvincing. The fact that we cannot stop driving and continue our heed-paying merely shows that we cannot continue to pay heed to something that is no longer there to pay heed to. We do not normally lapse into unconsciousness after applying the hand-brake; we turn our attention to other things. On the other hand the fact that one can, if one is sufficiently foolhardy, continue to drive and cease to pay heed to what one is doing would suggest *prima facie* that there are two distinct processes going on here. Ryle is doubtless right in pointing out that in driving with care one is not doing two things at once in quite the same sense as one is when one is walking along and humming at the same time. In humming and walking at the same time one is performing two distinct sets of muscular movements simultaneously. When heeding and driving occur together, on the other hand, there is only one set of muscular movements, those of manipulating the controls; and in that sense there is only one activity being performed. No one, however, supposes that heed-paying is a separate set of muscular movements occurring alongside the muscular movements involved in driving. Nor is heeding thought of as an unrelated activity going on at the same time as the driving. It is a peculiar sort of internal activity which controls the movements of the driver's limbs, by regulating his consciousness of the stimuli to which he responds.

V. THE DISPOSITIONAL THEORY OF MENTAL CONCEPTS

Although Ryle has failed to produce any conclusive objections to the contemplative theory, it is clear that if he can give a plausible account of the logic of 'heed concepts' which dispenses with the assumption that they refer to peculiar private events within the individual, we should undoubtedly be led to prefer such a theory on the grounds of parsimony. We must therefore examine the account which Ryle has offered of the logic of these concepts in order to discover whether or not it constitutes a satisfactory alternative to the traditional view.

The peculiarity of mental concepts as a class is that in order to determine whether or not someone knows, believes, understands, recognizes, remembers, wants, feels, is enjoying, attending to or thinking about something, you either have to cross-examine him or else observe considerable stretches of his behaviour before you can settle the question with any degree of confidence. This logical peculiarity is traditionally explained on the assumption that these mental concepts refer to invisible states and processes within the individual, whose existence and nature can only be determined with certainty by the individual in whom they occur, although it is usually possible also for an external observer to make reliable inferences about them by observing the behaviour to which they give rise.

Ryle's explanation of this logical feature is quite different. He supposes that mental concepts, or at least most of them, refer to what may be called behavioural dispositions, i.e. capacities, tendencies or temporary dispositions to behave in a certain way. To assert that someone has a capacity or tendency to behave in a certain way on this view is not to say anything about what is going on here and now, it is to assert a hypothetical proposition about how the individual could or would behave if certain circumstances were to arise. Hypothetical propositions of this kind can only be verified by investigating the behaviour of the individual under the conditions supposed. The proposition '*X* can swim', for example, can only be verified by observing *X*'s behaviour when in the water. Similarly with the proposition '*X* knows the date of the battle of Salamis': unless you happen at that moment to hear *X* say 'Salamis was fought in 480 B.C.', you would either have to wait for the chance of hearing his reactions when called upon to exhibit his knowledge of ancient history, or else adopt the more practical course of testing his knowledge by asking the appropriate question. The reason why it is often necessary to cross-examine the individual in order to discover what he knows, is not that knowing is a peculiar internal state or activity of the individual of which he alone is directly apprised; it is that an important part of what we mean when we say that *X* knows the date of the battle of Salamis is that he can give you the correct date when asked to do so.

With a few notable exceptions of which 'cogitating', 'visualizing' and 'having sensations' are the most important, Ryle attempts to apply this type of explanation to all the mental concepts treated in his book. In most cases moreover the attempt has proved remarkably successful. To my way of thinking there can be little doubt that the dispositional account which he gives of such concepts as 'knowing', 'believing', 'understanding', 'recognizing', 'remembering', 'intending' and 'wanting' is substantially correct. It is only with his attempt to apply it to such concepts as 'attending to', 'observing' and 'being conscious of something' that I wish to quarrel.

VI. RYLE'S APPLICATION OF THE DISPOSITIONAL THEORY TO HEED CONCEPTS

Ryle contends (pp. 137–9) that to say that someone is paying attention to what he is doing entails that he has at least two important dispositions, (*a*) the disposition under favourable circumstances to remember and give a first hand report on what it is he has been paying heed to, and (*b*) the disposition to adapt his performance to the various demands of the task as they arise. Now it is quite true that if we are told that someone is paying close attention to what he is doing, we normally expect him to be able to answer questions about his activity and to have made at least a better showing at the activity than if he had not been applying his mind to the same extent. It is also true, as Ryle points out, that we frequently conclude from the fact that someone is unable to answer questions about something that has been said in his presence, or from his failure in certain skilled performances, that he has not been paying attention to what was said or to what he was doing. But it does not follow from this that to say that someone is paying attention entails that he has the disposition to do these things. A schoolmaster frequently concludes from the fact that a boy has got the wrong answer to a mathematical problem, that he has set about it in the wrong way. Yet this would not lead us to say that to set about a problem in the right way entails a disposition to get the right answer. We only conclude that the boy must have used the wrong method if we know that his capacities are such that he could not have avoided getting the right answer had he used the correct method. Similarly, we only attribute someone's failure in a skilled activity to lack of attention if we know that his capacities are such that no other explanation of his failure is possible. One would hardly expect someone who had never been near an aeroplane before to be able to meet the demands of the task of piloting one, however closely he attended to what he was doing.

On the view which I am urging, the individual who pays attention is more likely to succeed in so far as he becomes acutely conscious of those features of the situation which are relevant to the successful performance of the task. Close attention to his own activity will be of no avail to the unskilled person because he has not learnt to discriminate between the relevant and irrelevant features. On the other hand an acute consciousness of the details of his own activity in relation to the environment may actually detract from the efficiency of performance in the case of an individual who has learnt to make many of the adjustments involved automatically. Thus we frequently say of someone whose skill is already well developed that his performance suffered because he paid too close attention to what he was doing. It is difficult to see what meaning could possibly be attached to this statement on a dispositional theory of attention.

In claiming that 'attending' entails 'being able to say something about what is going on', Ryle is on stronger ground. It is certainly true that it would be extremely odd to say that someone was paying attention to something, but could tell you absolutely nothing about it. But it is arguable that this is merely because one cannot pay attention to something without at least noticing the thing to which one is paying attention. There is no doubt that to say one has noticed something entails that one has the capacity to mention it and point it out, but to say that one has noticed something and to say that one has paid attention to it, are not, as Ryle appears to think, to say the same thing. 'Noticing' is an achievement concept like 'recognizing', 'perceiving', not an activity concept like 'pondering' or 'attending'. 'Noticing', 'perceiving', or 'recognizing', are the achievements that result from the activities of looking, listening and attending. If one looks, listens or attends one normally notices or recognizes something or other, but one can also attend and fail to notice; one can look and fail to see, listen and fail to hear. When we say that we have failed to notice anything, what we really mean is that we have failed to notice anything remarkable or anything additional to what we have already noticed. You can hardly be said to have paid attention, if there was nothing at all which you could be said to have noticed as a result of your attending. But it does not follow from the fact that *A* notices more about the situation than *B* that *A* was paying closer attention. The man who pays closer attention usually notices more, but the relationship is contingent rather than necessary.

VII. RYLE'S ACCOUNT OF THE LOGIC OF HEED CONCEPTS

The expression 'paying attention to something' exhibits the distinctive logical characteristics which are normally associated with words and expressions which refer to activities. The fact that it is perfectly good sense to speak of someone being engaged in paying attention to something, while it is nonsensical, for example, to speak of someone being engaged in knowing or understanding something, clearly shows that 'paying attention' or 'heed' is an activity expression in contrast to dispositional verbs like 'expect', 'know', 'like', and 'believe' or achievement verbs like 'understand', 'remember', 'recognize', 'perceive' and 'infer' where such a combination would be nonsensical.

It might be objected that 'attending' differs from those verbs which unquestionably refer to activities in that it is not sensible to use it in conjunction with adverbs like 'quickly' or 'slowly'. We can say 'he slowly began to pay attention to his surroundings', but not 'he paid slow attention to his book for five minutes and then rapid attention to the black-board'. There are, however, a number of expressions which can properly be described as 'activity verbs' of which the same is true. For example we can say 'he slowly took hold of the hammer' but not 'he held the hammer slowly for five minutes'. The analogy between 'attending' and 'holding' seems generally very close.

In support of his contention that 'attending' is not an ordinary activity verb, Ryle draws attention to the curious fact that it is always possible to replace a 'heed verb' by a 'heed adverb'. We can speak, to use his examples, of 'reading attentively', 'driving carefully' and 'conning studiously' just as readily as we can of 'attending to the page in front of one', 'taking care in one's driving', and 'applying one's mind to the task of translation' (p. 138). Ryle contends that the adverbial form is the more accurate way of expressing what is meant when a 'heed concept' is used. To say that someone is doing something heedfully, he maintains is merely to say that he is doing it in a certain way or in a certain frame of mind, i.e. with a disposition to adapt his performance to the various demands of the task as they arise and to answer questions about it.

On this theory the fact that 'paying attention' behaves like an ordinary activity word is explained on the assumption that the phrase 'paying attention to something' is analysable into two parts, (*a*) a categorical statement that a certain activity is taking place, and (*b*) a hypothetical or dispositional statement about how the individual in question would behave if certain contingencies were to arise. Ryle calls it for this reason a 'mongrel categorical statement'. The categorical part of the statement from which it derives its logical characteristics, is however, extremely uninformative. It asserts merely that some unspecified activity is being performed. In order to discover the nature of this activity we must find out what it is that the individual in question is paying attention to. As Ryle points out (p. 143) to say that someone is paying attention is an incomplete statement, unless we are told or unless it is obvious from the context of the remark what it is that he is paying attention to. On this view the part of the supposed meaning of the phrase which refers to the performance of an activity is strictly speaking redundant, since it must always be supplemented by a specification of the activity which is being performed.

In his lengthy discussion of 'mongrel categorical expressions' (pp. 140–7) Ryle is at pains to try to explain how it is that the fact of attention or inattention can be used to explain the failure or success of the individual in the activity he is performing. He shows convincingly enough that we can explain the bird's flying south by saying that it is migrating, without implying that migration is an additional process superimposed on the activity of flying south. The statement that the bird is migrating explains the behaviour of the bird by bringing the particular behaviour in question under the general rule that birds of certain species change their habitats at certain times of the year. Unfortunately he does not explain how the analogy is to be applied in the case where we explain the failure of an individual to complete a task satisfactorily or to give an adequate report on what was happening, by saying that he was not paying sufficient attention. It is difficult in this case to see what the general rule involved could be, unless it is the rule that if you don't pay attention you won't be able to carry out the activity you are performing satisfactorily or give an adequate first hand report on what went on. On Ryle's analysis of attention this general proposition reduces to the tautology 'unless you are disposed to give a first hand report on what is going on and to carry out what you are doing satisfactorily, you won't be able to give a first hand report on what is going on or carry out the activity you are performing satisfactorily'.

VIII. THE OBJECTION TO RYLE'S ACCOUNT OF THE LOGIC OF HEED CONCEPTS

Although Ryle fails to produce any conclusive reasons for adopting his theory of 'attending' in preference to the traditional account, it is difficult to produce any decisive arguments against it as long as we restrict the discussion to the special case where we speak of someone paying attention to what he is doing. His case breaks down, however, once we try to apply it to those cases where we are said to pay heed to an object in our environment or to some feeling we have, without being engaged in any other activity with respect to it. Ryle castigates the traditional theorists for 'misdescribing heed in the contemplative idiom' (p. 137), but he himself overlooks in developing his own theory the important cases where paying heed to something is purely a matter of watching, listening, observing or contemplating. Ryle explains the fact that 'attending' exhibits the usual characteristics of an activity verb, rather than those of a dispositional verb, on the assumption that verbs like 'attending' and 'heeding' assert the occurrence of the activities which are being performed attentively. There is no special activity called 'attending', there is only the attentive performance of an activity. The logical consequence of this theory is that the individual's own activities are the only sorts of things to which attention can be paid. If Ryle's theory were correct it should be nonsensical to talk of someone paying attention to anything other than an activity which he himself is performing. In fact, of course, we can speak with perfect propriety of the paying attention to any kind of object, phenomenon or sensation which is visible, audible, tangible, or otherwise perceptible. In such cases there is no activity which is being performed attentively or heedfully. To attend in such cases is merely a matter of contemplating or observing the object or phenomenon in question. We cannot say that when we pay heed to something we are watching it, listening to it, observing or contemplating it heedfully, since as Ryle himself points out (pp. 207 and 223), words such as 'watching', 'listening', 'observing' and 'descrying' already entail that heed is being paid. These expressions do not refer to activities like driving a car which can be performed with or without heed, they refer to special forms of the activity of heed-paying itself. It makes nonsense to say that someone was observing, watching, contemplating or listening to something without paying any attention to it, whereas it makes perfect sense to speak of someone driving without paying any attention to what he is doing.

The inadequacy of Ryle's account appears most clearly when we examine the account which he gives of expressions like 'being conscious of', 'observing', 'watching' and 'listening'. To be conscious of the sensations in a blistered heel according to Ryle (pp. 157–8) is to pay heed to them; but what is the activity which is being performed attentively or heedfully here? It would seem from his long discussion of 'observation' (ch. VII) that for Ryle to say that one is observing something is to say that one is paying heed to the sensations derived from it, and 'watching' and 'listening' by the same token, refer to the paying of heed to visual and auditory sensations respectively. But we cannot say that 'having sensations' is the activity which is being performed heedfully in these cases, since to have a sensation itself entails paying at least some heed to the sensation. We can speak of failing to notice the sensations which one would have had if one had paid attention to them; but to say that one had a tingling sensation in the left toe without noticing it is nonsense. Ryle accuses the traditional theory of being unable to provide a sensible account of the difference between a careful and a careless observer, but his own theory, while giving a plausible account of carefulness, fails to explain the activity of observing.

IX. THE DISPOSITIONAL THEORY RESTATED

Although Ryle has failed to provide a satisfactory account of consciousness, attention and observation in terms of dispositional theory of mental concepts, it would be unwise to conclude that such an account cannot be given. Ryle's account fails mainly because he overlooks the fact that our own activity is not the only sort of thing to which we can pay attention. The possibility of providing a plausible dispositional theory which takes account of our consciousness of and attention to objects, phenomena and sensations is not ruled out. Indeed it is not difficult to suggest the form which such a theory might take.

We have seen that although paying attention to what one is doing does not entail being prepared to meet the demands of the task in hand, it cannot be denied that to pay attention to something entails noticing and hence being able to say something about it. It must also, I think, be conceded that it involves being ready to encounter something, although one need not be prepared to encounter the sort of thing that is actually there. This, however, cannot be all that we are saying when we say that someone is attending to something, since we can be ready to behave in a manner appropriate to the presence of some object or event in our immediate environment without actually being conscious of it. Our disposition to act in this way may be a result of something we have been told, some inference we have drawn or some observation made a few moments previously. In such cases we might be said to know, remember or suspect that it was there, but we would not be observing, attending to or conscious of it. In order for us to be conscious of something our disposition to react to its presence must result from its impingement on our sense organs at the time.

With the qualification that the disposition must result from sensory stimulation, it becomes quite plausible to maintain that to be conscious of something is to be ready to react both verbally and otherwise to the presence of some object or event in one's immediate environment. On this theory the contribution of attending to skilled performance would be explained by pointing out that unless the individual is disposed to react in a manner appropriate to the presence of the relevant features of his own activity and the environmental situation in which it takes place, he is not likely to be very successful.

On this view 'consciousness', 'attention' and 'observation' refer to a temporary state of readiness for something. You would therefore expect them to exhibit the logical features of expressions referring to temporary states of affairs. Expressions like 'being conscious' or 'aware of' do exhibit these logical characteristics. Words like 'attending', 'observing', 'watching', 'looking' and 'listening', on the other hand, exhibit the logical behaviour characteristic of expressions which refer to activities. This fact as we have seen, appears to provide a formidable obstacle to any dispositional theory of the meaning of these words. Nevertheless the difficulty can probably be overcome without appealing to any kind of internal process or activity, by examining the notion of 'activity' itself. It is at least arguable that when we speak of an individual's activities, of the things he does, we refer to those changes in him which can be induced by such things as commands, entreaties, instructions and deliberations. Any changes whether muscular or non-muscular which he can decide or be asked to bring about in himself are things which he does. Paying attention and observing are not muscular movements, nor are they movements of a mysterious transcendental musculature, they are, so the theory might run, changes in the individual's short term dispositions, readinesses or sets (to use a term which has a wide currency in the psychological literature) which can be induced by appropriate commands, requests or by decisions on the part of the individual himself.

In the light of these considerations we may restate the dispositional theory of attention, observation and consciousness as follows: to observe or pay attention to something is to bring about a change in oneself such that the impingement of the object or phenomenon in question on one's receptor organs prepares one to respond both verbally and otherwise in a manner appropriate to the presence of something; while to be conscious of something is to be so disposed.

X. THE CASE FOR THE TRADITIONAL VIEW

Stated in this way my quarrel with the dispositional theory is less substantial than my agreement with it. My contention is not so much that it is wrong as that it is incomplete. It is incomplete because it makes no reference to the internal state of the individual which enables him to describe and respond appropriately to the presence of objects in his vicinity. On the view which I wish to defend, when we use what Ryle calls a 'heed concept', we are not merely referring to the disposition to respond in a manner appropriate to the presence of the thing in question and specifying how that disposition is brought into being, we are also referring to an internal state of the individual which is a necessary and sufficient condition of the presence of such a disposition. I shall now try to present arguments in support of this contention.

One of the major weaknesses of Ryle's account of mental concepts is, as he himself recognizes, his retention of the traditional extended use of the term sensation (ch. VII). He is compelled to retain this use in order to provide something—having a sensation—which the observer of an object can be said to do heedfully. One of the advantages of the revised form of the dispositional theory which I have stated is that it dispenses with the necessity for this concession to the traditional misappropriation of mental concepts. But although it dispenses with the necessity of abusing the concept of sensation, it runs into serious difficulties when applied to those cases where we do speak of being conscious of or of attending to our sensations, i.e. in those cases where our state of consciousness results from interoceptive or proprioceptive stimulation or from the various twists and quirks of our sensory apparatus, rather than from the impingement on our sense organs of any specifiable state of affairs in our environment.

Suppose that having applied pressure to my eyeball, I am conscious of a sensation of light. According to the revised dispositional theory this means that I am disposed to react to the presence of something. But what is it that I am disposed to react to? It cannot be the pressure on my eyeball, since to be conscious of light sensations is not the same thing as being conscious of pressure applied to the eyeball. But it cannot be the presence of the sensation either. Sensations, as we have seen, do not exist independently of our consciousness of them. There are not two things, my sensation and my consciousness of it, in the way that there are two things, a penny and my consciousness of the penny. The occurrence of a sensation entails someone's consciousness of that sensation.

To be disposed to react to a sensation therefore would be to be disposed to react to one's consciousness of that sensation. In other words we now have an infinite regress of dispositions, instead of the infinite regress of ghostly operations which appears so frequently in Ryle's criticisms of the traditional theories. We might be tempted to meet this objection by supposing that to say that someone is conscious of a sensation of light is to say that he is temporarily disposed to react as he would normally do if there had been a flash of light. But to be disposed to react as if there were a flash of light, would be to believe or be tempted to believe that a flash had occurred; whereas it makes perfectly good sense to say that he was conscious of a vivid sensation of light, yet it never occurred to him for one moment to suppose that there had been any actual flash of light. In other words an individual's state of consciousness is something over and above any dispositions which it arouses in him.

An objection which applies to any attempt to give a dispositional amount of consciousness and attention is the objection that it always makes sense to ask the individual to describe what it is like to watch, listen, observe or be conscious of something, whereas it does not make sense to ask him what it is like to have a certain capacity or tendency. We can only describe what something is like if it is an object, situation or occurrence. We can describe, characterize or define such things as relationships, capacities and tendencies, but we cannot describe what they are like. We can describe what a car is like, but we cannot describe what its horse-power is like; we can describe what it is like for one billiard ball to strike another and propel it forward, but we cannot describe what the causal relationship is like; we can describe what it is like to swim or what it is like to realize that one can swim, but not what it is like to be able to swim; we may be able to describe what it is like to be told or call to mind the fact that whales are mammals, but we cannot describe what it is like to know or believe that they are. If to be conscious of something were merely to be disposed to react in some way, it should be logically impossible for us to describe what it is like to be conscious of something. In fact there is no logical impossibility here. We are continually describing what it is like to watch, look at, listen to or feel things.

It might be objected with some justification here that what we describe is not our consciousness, but the things we are conscious of. As we have seen, part of what is meant by saying that someone is conscious of something is that he can say something about it. It is certainly true that when we describe some object in our environment of which we are conscious, our description is a description of the object itself, and not, as has sometimes been supposed, a description of our consciousness of that object. It is also true that we cannot describe the state of being conscious in abstraction from the things we are conscious of. But that does not mean that we do not on occasions describe our consciousness of things as distinct from describing the things themselves. When we say, to use a familiar example, that the penny looks elliptical when viewed at an angle, we are not describing the penny, nor are we describing the image which it projects on our retina; we are describing what it is like to look at a penny from that particular angle; we are saying that it is somehow like looking at an ellipse viewed full face. When we say this, moreover, we do not imply that we are disposed to act in a manner appropriate to its being an ellipse. The elliptical shape of the penny is not an optical or a psychological illusion (cf. Ryle's discussion of this problem pp. 216–18).

When we describe a state of consciousness, we usually do so by comparing being conscious of one thing with being conscious of another. Nevertheless there are one or two expressions like 'pleasant', 'unpleasant', 'vivid', 'dim', 'acute' and 'vague' which we apply to the states of consciousness themselves. These are somewhat unusual adjectives to apply to a state of readiness. Furthermore the difference between vividness and dimness, acuteness and vagueness is difficult to explain on a dispositional theory of consciousness. The only possible interpretation on such a theory is in terms of the appropriateness of the behaviour, for which one is prepared, to the presence of whatever it is one is conscious of. Acute consciousness, however, does not guarantee the appropriateness of the resulting behaviour. The statement 'his consciousness of his own ineptitude was so acute that he was unable to do anything about it', makes perfectly good sense. It also describes a situation with which some of us are only too painfully familiar. If we recognize that consciousness is some sort of internal state of the individual these discrepancies between the intensity of the individual's consciousness and the adequacy of the behaviour for which it prepares him, no longer constitute a problem.

Finally, there are considerations of a more general nature. If there were no decisive arguments either way, we should probably prefer the dispositional to the internal process theory of consciousness and attention on the grounds of parsimony. As against this must be set the fact that in every other case where verbs having 'activity' characteristics are involved, it has been found impossible to apply a purely dispositional analysis, and in at least one group of cases the reference to internal processes within the individual cannot seriously be denied. The cases I have in mind here are thinking (in the sense of thinking about or thinking to oneself), pondering, calculating, imagining, dreaming, visualizing and doing mental arithmetic. Ryle (p. 27) has made a strong case for the view, that when we talk about someone thinking (in the relevant sense), pondering, calculating, or imagining we are not asserting the occurrence of any internal process or activity. He contends that the activity referred to, although sometimes covert, as when it consists of visualizing or performing mental arithmetic, need not be so. It may equally well consist in some entirely overt performance such as drawing, talking out loud to oneself or playing a game of make-believe. To assert that someone is thinking or imagining does not discriminate between these two possibilities. This argument disposes or, at least appears to dispose of, the view that words like 'thinking' and 'imagining' necessarily assert the occurrence of covert activities, but there is no suggestion that these are dispositional concepts. Nor is there any attempt to deny that thinking sometimes consists in a purely covert process or that expressions like 'dreaming', 'visualizing' and 'mental arithmetic' refer to such processes. If this is conceded with respect to 'dreaming', 'visualizing' and 'mental arithmetic', it is difficult in view of the weight of traditional and common-sense opinion and the lack of any positive evidence against it, to see why a similar concession should not be made with respect to 'attending', 'observing', 'watching', 'looking' and 'listening'.

Concepts like 'observing', 'watching', 'listening' and 'being conscious of' are, in fact, closely related to concepts like visualizing and dreaming in a way, which is extremely difficult to explain, if the former are regarded as dispositional concepts. For if we want to explain what sort of thing this business of visualizing or dreaming is, the answer which immediately suggests itself is to say that visualizing something is like watching it except that there is nothing there really and you don't have to have your eyes open. Now if to watch something is merely to bring about a change in oneself such that the impingement of the thing in question on one's eyes prepares one to respond both verbally and otherwise in a manner appropriate to there being something there, this explanation becomes completely unintelligible. Apart from the fact that both visualizing and watching are things which the individual can be said to do, it is exceedingly difficult on this theory to find anything which the two cases have in common. We cannot say that to visualize is to be disposed to act and speak as if there were something impinging on one's eyes when in fact there is not. Any one who is so disposed would be suffering from a visual hallucination, and although having a visual hallucination may be said to involve visualizing, we can visualize things perfectly well without being hallucinated. The similarity between visualizing something and watching it lies in the internal state of the individual which is brought into being, not in the behavioural dispositions which that state induces.

XI. CONCLUSIONS

If the above arguments prove what I think they prove, are we back where we started at the beginning of Ryle's inquiry? Do these arguments merely put the Ghost back into the Machine? I do not think so. So far as I am aware, the criticisms I have made of the dispositional theory apply only to the dispositional analysis of consciousness and heed concepts generally. The dispositional analysis of intelligence, knowledge, belief, motives and memory remains unaffected, except in so far as these concepts involve dispositions to pay attention to or become conscious of certain features of one's environment. Indeed, since Ryle himself appears to accept the view that words like 'watching', 'listening' and 'observing' entail a reference to a covert process of having sensations, it is only in the case of the heedful performance of muscular activities that the view which has been urged in this paper differs from the account which Ryle has given as far as recognizing a reference to covert states and processes is concerned. On Ryle's view, however, these processes are relatively unimportant; we learn to talk silently to ourselves in order not to disturb others; we could plan our course of action on paper, but it is often more convenient to do it in our heads. If, on the other hand, our very ability to describe and adapt our behaviour to the objects and phenomena which impinge on our sense organs, is dependent on a special state of affairs within ourselves, which can itself be described by the person in whom it occurs, the reference which is made to such a process in our use of expressions like 'attending', 'observing' and 'being conscious' can hardly be brushed aside as a matter of no great significance. If such a view is accepted, we can hardly avoid raising the question which Ryle has dodged persistently throughout his book, namely the question: 'What are these curious occurrences within ourselves on which we can give a running commentary as they occur?' Lack of space unfortunately precludes any discussion of this fascinating problem here. It is my belief, however, that the logical objections to the statement 'consciousness is a process in the brain' are no greater than the logical objections which might be raised to the statement 'lightning is a motion of electric charges'.

REFERENCE

RYLE, G. (1949). *The Concept of Mind.* London, Hutchinson's University Library.

(*Manuscript received* 23 *February* 1954)

AN EXPERIMENTAL STUDY OF HUMAN CURIOSITY

By D. E. BERLYNE

University of Aberdeen

I. *Introduction* (p. 256). II. *Predictions from the theory* (pp. 256–259). III. *The experiment* (pp. 259–260). IV. *Results and discussion* (pp. 260–265). V. *Summary* (p. 265). *References* (p. 265).

I. INTRODUCTION

A drive which is aroused by a question and reduced by rehearsing its answer is identifiable with what we have called 'epistemic curiosity' (Berlyne, 1954). The present paper reports an experiment intended to test some of the predictions derivable from the theory that was built round this concept.

When we set out to inquire into a complex form of human motivation like curiosity, we find ourselves faced with a bewildering array of variables that may be relevant. The difficulty of knowing where to begin is, no doubt, one reason why little work in this area has been done. One indispensable aid is to have a theory (Hebb, 1951; Miller, 1951) to suggest relationships that are likely to repay investigation. But even then, the task may still seem baffling. There may be a vast network of factors involved, each making a comparatively slight contribution, and individual differences must be enormous. It seems therefore desirable to pass through an intermediate stage, if the project is to be practicable, namely an exploratory experiment. This would use small samples of subjects and sound the effects of several variables at once. It may well prove too insensitive to permit definitive conclusions about some of the relationships it studies. But it can save us from many a costly blind alley by confirming that certain lines of research are worth pursuing. The experiment to be reported below was designed for such an exploratory function.

II. PREDICTIONS FROM THE THEORY

The theory with which we are concerned (Berlyne, 1954) describes certain postulated, unobservable central processes. It is therefore necessary to relate our intervening variables to antecedent and consequent variables, in order that the empirical content of the theory may be revealed.

Epistemic curiosity, according to the theory, is aroused by 'thematic probes' (Skinner, 1947, 1953). The clearest examples of these are questions, which are assumed to evoke, among others, drive-producing meaning-responses (r_{mD}). The curiosity-drive-strength is assumed to increase with (1) the intensity of the drive-stimulus (s_{mD}) produced by the r_{mD}, and (2) the degree of conflict (F) between the symbolic meaning-responses.

The intensity of the s_{mD} will depend on the prolongation (extinction-resistance, n), and the strength (amplitude, A) of the drive-producing response (r_{mD}). These quantities, according to Hull (1943, 1952), depend on reaction-potential (${}_{S}E_{R}$), which is itself largely a function of habit-strength (${}_{S}H_{R}$), inhibitory aggregate ($\dot{I}_{R}$) and drive (D). To take these in turn, ${}_{S}H_{R}$ will be a function of the number of reinforcements (N), which in this case will mean the number of times that arousal of curiosity by similar questions in the past has led to a successful quest for an answer, with consequent reward, social or material.

Then, $\dot{I}_R$ will depend on the difficulty of finding an answer to similar questions in the past (amount of effort, W) and the number of times the quest has led to failure (n). As for D, Hull postulates that any drives present at the time will help to swell this variable, but in our case it is likely that irrelevant drives like hunger or thirst will evoke their own distracting responses more strongly than they will increase curiosity. But drives whose relief has on previous occasions been furthered by finding answers to comparable questions may well add their quota. They may include practical purposes for which the knowledge sought would be of use, social motives, e.g. drives for prestige and approval, and possibly such drives as Freudians mention, acting by displacement (Miller, 1948). The other variables on which ${}_SE_R$ is assumed to depend are likely to play a minor role in curiosity, but they may possibly not be without influence. They are J, or the speed with which reward has followed a successful inquiry in the past (this is probably of little importance because secondary rewards or 'confirmatory reactions' are likely to be immediate) and V (stimulus-intensity dynamism), which may represent the effect of the loudness, if spoken, or size and brightness, if written, of the question.

But the more innovative aspect of the theory, and the one with which the experiment is most closely concerned, is that which deals with the role of conflict (F). A discussion of the general design of the experiment will show how its effects may be studied empirically.

An experimental group received a fore-questionnaire about invertebrate animals, a communication providing answers to the questions and an after-questionnaire repeating the questions. A control group was given the communication and the after-questionnaire but not the fore-questionnaire. In this situation the theory would predict several phenomena:

(1) *Effects of pre-questioning*

It would be expected that the experimental group, when they read one of the answers, would recall the corresponding question from the fore-questionnaire, and that curiosity would be re-aroused as a result. The rehearsal of the answer would reduce the curiosity-drive to a subliminal value, and this drive-reduction would reinforce the learning of the answer. No such process could occur with the control group, as they had no fore-questionnaire. Other sources of reinforcement might well be present, but the experimental group, having this extra drive-reduction, would learn more effectively. They would therefore recall more answers than the control group when the after-questionnaire was presented.

The arousal and subsequent reduction of curiosity is stated to increase, by stimulus generalization, the amplitude of curiosity-producing responses (r_{mD}) to similar questions in future. Consequently, when subjects are required to indicate which of the animals figuring in the experiment they would like to know more about, the experimental group can be expected to mark more animals than the control group.

(2) *Recognition of answers*

When one of the answers in the communication is received, curiosity aroused by the recall of the question will be reduced. If a question is recalled, its answer will presumably be recognized as such. It follows that the learning of statements which are recognized as answers to questions from the fore-questionnaire will have the reinforcement of this curiosity-reduction, and so they will be more probably recalled than others during the after-questionnaire.

(3) *Effectiveness of different classes of questions*

We have two ways of comparing the effectiveness in arousing curiosity of two different classes of questions. Subjects were instructed to mark with 'K' twelve questions out of the forty-eight in the fore-questionnaire whose answers they most wanted to know, and so we can note the number of questions in each class that were marked 'K'. Secondly, we can note the number of questions in each class that were answered correctly during the after-questionnaire.

The conflict-variable (F), which the theory cites as a determinant of strength of curiosity, is stated to depend on the equality, the absolute strength, the number and the degree of incompatibility of the competing response-tendencies. This allows us to identify certain classes of questions which can be expected to arouse more curiosity than others, as shown by both ways of gauging curiosity.

Familiarity. It is hypothesized that the more familiar the principal concepts in the questions, the more numerous and the stronger the competing symbolic response-sequences to which they give rise. Noble (1952, 1953) has demonstrated that there is a correlation between the familiarity of words and the number of associations they elicit. Subjects were instructed to rate the animals figuring in the questions for familiarity at the start of the experiment, and familiarity was also manipulated by providing introductory information about two of the animals. Questions about more familiar animals, familiarity being judged by either operation, should arouse more curiosity than others.

Incompatibility. Questions whose concepts seem incompatible to subjects will arouse more conflict than others and consequently more curiosity. Incompatibility was judged by having subjects mark those questions in the fore-questionnaire which surprised them, and also by using a group of judges, who indicated which predicates seemed to them least applicable to the animals.

A preliminary experiment, reported elsewhere (Berlyne, 1953, appendix), supported the prediction that questions with four alternative answers aroused more curiosity than those with two alternative answers, as we should expect in view of the number-of-competing-tendencies variable.

(4) *Correlation between measures of curiosity*

The questions which subjects marked with 'K', as arousing most curiosity in the fore-questionnaire, should be more likely than others to be answered correctly in the after-questionnaire. There are two mechanisms by which this could occur. Whether a question is marked 'K' depends on how much curiosity it arouses in the fore-questionnaire, and whether it is answered correctly in the after-questionnaire is assumed to depend on the strength of curiosity at the time of reading the answer in the communication. The theory implies that these two degrees of curiosity will be correlated. A supplementary or alternative mechanism might be a closer reading of questions selected for marking with 'K' and consequently a more probable recall of such questions in the communication. Either mechanism would cause our two measures to be correlated and would make both depend on the degree of curiosity aroused by the questions.

(5) *Surprisingness of answers*

Subjects were instructed to indicate which statements in the communication they found surprising. Such statements will have some apparent incompatibility between their concepts, and this will add to the curiosity aroused and then reduced. These statements should therefore be more effectively learned and more probably recalled in the after-questionnaire than others.

There were sixty-eight subjects and thirty judges; all were undergraduates taking a course in Normal Human Personality at Brooklyn College. All had had a previous introductory course in General Psychology. Subjects occupying alternate seats were allotted to the experimental group and to the control group.

III. THE EXPERIMENT

SUBJECTS. *Experimental Group* (thirty-four subjects).

The following three phases were taken in immediate succession.

Phase I:

Section 1. A list of twelve animals was presented, and subjects were instructed to rate each animal for familiarity on a 4-point scale, ranging from 'seen, read or heard a fair amount of' to 'never heard of'. The twelve animals consisted of eight that subjects were almost bound to have heard of and four 'exotic' animals that were unlikely to be known to them. Of the 'exotic' animals, two were real (the sea-mat and the sea-gooseberry) and two were fictitious (the 'sea-wasp' and the 'stringworm').

Section 2. Paragraphs of about 120 words each, giving a general description of two of the 'exotic' animals, one real and one fictitious, were read by subject. For half the subjects (sub-group EA) the animals were the sea-gooseberry and the 'stringworm', while for the remainder (sub-group EB) they were the sea-mat and the 'sea-wasp'.

Section 3. A randomized series of 48 questions, 4 on each animal, was presented, and subject had (*a*) to check off the correct answers from two alternatives per question, (*b*) to mark with 'C' those answers of which he felt certain from previous knowledge, and (*c*) to mark with 'S' those questions which surprised him.

Section 4. Subject had to go back over the questions in section 3 and mark with 'K' the three questions out of each consecutive set of twelve whose answers he would most like to know. Twelve questions out of forty-eight were thus to be so marked.

Phase II:

Subject received sheets containing seventy-two statements, six about each of the twelve animals. These were in a random order, not corresponding to the random order of the questions in Phase I. Subject was (*a*) to mark with 'S' every statement that surprised him, and (*b*) to tick every statement that he recognized as an answer to one of the questions in Phase I. An answer to each of the forty-eight questions, coinciding with one of the alternatives offered in Phase I, was included among the seventy-two statements. The tasks served the double purpose of providing the experimenter with information and inducing, as far as possible, a non-learning set.

Phase III:

Section 1. The forty-eight critical questions were again presented, this time in an open-ended form and in yet a third random order. Subject had to fill in the answer given in Phase II, when remembered.

Section 2. The twelve animals were listed, and subject had to check off those he would like to know more about. Finally, there were questions on the interestingness of the experiment and desire to know more about animals.

Control Group (thirty-four subjects).

The procedure was exactly the same as for the experimental group, except that sections 3 and 4 of Phase I (the fore-questionnaire) were omitted, and the recognition-test in Phase II did not apply. Phase I existed in two versions, as with the experimental group: seventeen subjects (sub-group CA) received paragraphs about the sea-gooseberry and the 'stringworm', and the other seventeen (sub-group CB) read about the sea-mat and the 'sea-wasp'.

JUDGES (thirty).

The judges, taken from the same population as the subjects, received a list of the eight 'non-exotic' animals used in the experiment. Opposite the name of each animal were four phrases, representing the predicates implied by each of the four questions about that animal. The judges had to check off the two phrases that seemed to them least likely to apply to the animal concerned.

Of the forty-eight answers to questions from the fore-questionnaire that were presented in Phase II, twenty-four (two per animal) were true and twenty-four (two per animal) were false. This was to control for previous knowledge of the subject-matter and for the fact that the animals mentioned differed in familiarity. If all the answers provided had been true, more knowledgeable subjects would have had more coincidences between the answers checked in Phase I and those learned in Phase II, and the following possibilities might have invalidated the conclusions:

(1) Any previous knowledge that inclined a subject to prefer a particular response in Phase I may have made the corresponding fact easier to learn in Phase II.

(2) The act of selecting an answer in Phase I may have provided reinforcement for the bond between question and answer to supplement the learning in Phase II.

The familiarity-variable necessitated precautions against two further possibilities:

(1) The correct answers to questions about familiar animals might have been encountered in the past. These answers might have been subliminal, so that they could not be recalled in Phase I, but they might nevertheless have provided a residue of habit-strength to make the re-learning of them in Phase II easier.

(2) It might have been possible to infer the correct answer from what was already known about the animals. (We guarded against this also by the inclusion of fictitious animals, the answers being selected at random in their case.)

The expedient of making half the answers false ensured that, no matter how much a subject's prior knowledge about the animals, this knowledge would facilitate and hinder the learning of equal numbers of answers. It caused the proportion of coincidences per subject between answers checked in Phase I and answers learned in Phase II to approximate one-half.

IV. RESULTS AND DISCUSSION

Computing the results

The following rules were observed in computing the results, except where stated otherwise.

(1) All questions marked 'C' in the fore-questionnaire (i.e. those whose answers subjects claimed to know with certainty) were omitted. This was because in such cases curiosity may not be aroused at all or may function in a special way, and different subjects felt certain of different numbers of answers.

(2) In some cases, subjects appeared to have misunderstood or overlooked some of the tasks. Therefore, (*a*) if subject had marked no question with 'S' (to indicate surprise) in Phase I, his data were omitted from all calculations involving that variable, (*b*) if subject marked all statements in Phase II with either 'S' (indicating surprise) or with a tick (indicating recognition of an answer), but not both, his data were disregarded where these variables were relevant, (*c*) if subject marked more or less than twelve of the questions in Phase I with 'K' (indicating desire to know the answer) his data were included, but if he marked none at all they were disregarded where relevant.

For the comparisons between experimental and control groups, we used the *t*-test. But for other results, which involve comparisons between questions or answers, differing in number from subject to subject, that belonged to different categories, we used the χ^2-test for data pooled from all subjects, with Yates's correction. Some statisticians object to the use of the χ^2 for such data, one objection being that the test does not distinguish between effects which the group as a whole has produced and those due largely to a few exceptional individuals. We have consequently given, in columns headed 'Ss' in the tables, the number of subjects whose data show a deviation from expected proportions in the same direction as the contingency table for the whole group. For example, 26/29 indicates that twenty-six out of twenty-nine subjects show such a deviation. This enables us to supplement the χ^2 analysis with the sign-test (Moses, 1952). Our experiment, having mainly an exploratory function, is unfortunately not sensitive enough for such techniques as the analysis of variance for proportions based on unequal numbers (Cochran, 1943).

Effects of pre-questioning. Table 1 displays the comparisons between experimental and control group that show whether the pre-questioning, which only the experimental group underwent, increased curiosity. We compare the mean numbers of questions answered correctly in Phase III and also the mean numbers of animals ticked in section 2, Phase III. Both differences confirm the prediction significantly. All questions, including those marked 'C' are included. No significant differences were produced by any of the questions on interest, etc.

In view of the fact that the experimental group had two tasks to carry out in Phase II and the control group had only one, it might be thought that some difference in the thoroughness of reading the statements might have been responsible for the greater recall-scores of the experimental group. But a similar significant difference was found in the preliminary experiment (Berlyne, 1953, appendix), in which both groups alike had merely to read through the communication in Phase II.

Table 1. *Effects of pre-questioning*

Mean correct answers in phase III		*t*	*P*	Mean animals ticked in Phase III		*t*	*P*
Exp. group	Cont. group			Exp. group	Cont. group		
32·41	27·15	3·29	<0·01	5·38	3·36	2·90	<0·01

Recognition of answers. The first row of Table 2 shows that statements recognized by the experimental group as answers to questions from Phase I were, as predicted, more likely than others to be recalled in the after-questionnaire. The second row of the table shows that questions subjects marked 'K' (i.e. those whose answers they most wished to know) in Phase I were more likely to have their answers recognized in Phase II. The latter relationship was not predicted by the theory. To explain it, we can only hazard the

Table 2. *Relations between recognition of answer and measures of curiosity*

Questions	Answers recognized	Answers not recognized	χ^2	P	Ss
Answered correctly in Phase III	732	72	31·38	<0·0001	26/29**
Not answered correctly	325	91	—	—	—
Marked K	256	24	7·25	<0·01	21·5/27**
Not marked K	780	138	—	—	—

** Significant at 0·01 level by sign-test.

hypotheses that (1) questions evoking more curiosity may, like other drive-arousing stimuli, receive more attention (Berlyne, 1951, p. 143) and thus be more easily remembered, (2) questions marked 'K' may receive more attention, or (3) the drive-producing responses, though subliminal, may persist in some form, so as to supplement the statements in Phase II as stimuli re-evoking the questions. Such subliminal drives, known as 'quasi-needs' or 'tensions', have been posited by Lewin and his followers (1935) as an explanation of the Zeigarnik effect. They are also reminiscent of Freud's 'unconscious wishes', which may likewise lie dormant until they determine a response jointly with incoming external stimuli.

Effectiveness of different classes of questions. The last four rows of Table 3 present comparisons between different classes of questions put to the experimental group according to our two ways of estimating the effectiveness with which they arouse curiosity. We examine the number of questions in each class answered correctly and the number answered incorrectly in Phase III, and similarly we compare the numbers marked 'K' (as among the twelve whose answers subjects most wished to know) and not marked 'K' in Phase I.

Familiarity. We compare questions about animals marked 'seen, read or heard a fair amount' or 'seen, read or heard a little' (hereafter referred to as the '*more familiar animals*') with those marked 'heard of, but know nothing about' or 'never heard of' (hereafter referred to as the '*less familiar animals*'). Table 3 shows that questions about more familiar animals evoked more curiosity by both tests. The two 'exotic' animals

Table 3. *Effectiveness of different classes of questions*

(N.S. = 'not significant')

Classes of questions	Correct in Phase III	Not correct in Phase III	χ^2	P	Ss	Marked K	Not marked K	χ^2	P	Ss
Marked K	251	76	20·11	<0·0001	28/31**	—	—	—	—	—
Not marked K	670	390	—	—	—	—	—	—	—	—
More familiar animals	659	239	44·50	<0·0001	29/34**	259	535	36·70	<0·0001	26/32**
Less familiar animals	187	163	—	—	—	50	286	—	—	—
Previous-information 'exotic' animals	149	111	NS	—	—	28	213	NS	—	—
No-previous-information 'exotic' animals	138	135	—	—	—	23	225	—	—	—
Surprising	153	88	NS	—	—	81	130	20·11	<0·0001	23/28**
Non-surprising	751	411	—	—	—	230	761	—	—	—
Incompatible concepts	362	159	NS	—	—	169	293	6·98	<0·01	24·5/31**
Non-incompatible concepts	355	140	—	—	—	131	329	—	—	—

** Significant at 0·01 level by sign-test.

about which subjects received previous information were omitted from this comparison, as they received a special treatment. The other two were included, except where a fictitious animal received a rating other than 'never heard of'.

It might be suggested that the better recall-scores for questions about more familiar animals were due to the fact that more of these questions were marked 'K', which may have caused them to be read more attentively in Phase I and perhaps remembered more easily in Phase II. However, in the preliminary experiment (Berlyne, 1953, appendix), only half the subjects in the experimental group had to carry out the task of marking questions with 'K', and no significant effects on recall-scores of the presence or absence of this operation were found.

Table 4. *Recall of different classes of statements*

(NS = 'not significant')

Group	Classes of statements	Recalled in Phase III	Not recalled in Phase III	χ^2	P	Ss
Cont.	More familiar animals	498	336	24·50	<0·0001	26·5/34**
	Less familiar animals	178	223	—	—	—
Cont.	Previous-information 'exotic' animals	127	135	6·33	<0·05	22·5/34
	No-previous information 'exotic' animals	102	169	—	—	—
Cont.	Surprising	369	292	NS	—	—
	Non-surprising	438	392	—	—	—
Exp.	Surprising	188	75	5·06	<0·05	21/29*
	Non-surprising	662	375	—	—	—

* Significant at 0·05 level by sign-test.
** Significant at 0·01 level by sign-test.

The effects of *questions* about familiar animals must be distinguished from any tendency there might be to remember *facts* about familiar animals better. The data on the control group in Table 4 reveal that there is such a tendency. However, we can ascertain whether questions about more familiar animals have an effect over and above this by comparing for the two groups the proportions of *questions answered correctly* in Phase III that fell in the two categories. We already know that pre-questioning increases the probability of recalling facts. But we wish to know whether this increase is significantly greater for questions about more familiar animals, i.e. whether such questions arouse more curiosity. If they do not, we should expect the correctly answered questions to be distributed in equal proportions between more familiar and less familiar animals for both groups. Table 5 supports the prediction that more curiosity is aroused by questions relating to more familiar animals. In the case of the control group, the figures are adjusted, because those for the experimental group excluded questions marked 'C' (those whose answers subjects claimed to be certain of). It is assumed that the control group would have marked with 'C' the same proportions of questions in each category as the experimental group, and the totals are reduced by these proportions.

Each of the four '*exotic*' animals received an introductory paragraph of information for half the subjects of each group, but not for the other half. We are thus enabled, confining our attention to the 'exotic' animals alone, to compare the effects of questions about *previous-information* and *non-previous information* animals respectively. Table 3 shows that there is no significant difference in curiosity by either test. But the data from the control group in Table 4 (adjusted, as above) suggest there is a tendency for facts about previous-information animals to be learned better. The failure of a significant difference to appear for the experimental group may well have been due to its being swamped by the pre-questioning variable, which may have produced a ceiling-effect.

Table 5. *Questions answered correctly in Phase III*

Group	More familiar animals	Less familiar animals	χ^2	P
Exp.	659	187	3·46	<0.05 (1 tail)
Cont.	498	178	—	—

Incompatibility. There are two indications available to us of which questions contradicted subjects' expectations. We have the marking of questions with 'S' in Phase I, which reveals those found *surprising*, and we have the judges' selection of the two questions out of the four on each animal that involved what appeared to them the least applicable predicates. The latter dichotomy between '*incompatible-concept*' and '*non-incompatible-concept*' questions has the advantage of being orthogonal to any comparisons between animals on the grounds of familiarity, etc., as each animal had two questions in each category. Table 3 shows that, according subjects' reports (marking 'K') of desire to know answers, surprising and incompatible-concept questions evoked more curiosity than the others. But no effect of these variables on recall was manifested. We can check on the agreement among the judges in their selection of incompatible concepts by counting the number of judges marking the two most and the number marking the two least frequently selected concepts for each animal, and performing a mean-difference test. The mean difference per animal is 14·52, $t = 4·09$ and P is less than 0·01. The preliminary experiment (Berlyne, 1953, appendix) supplied some slight evidence in favour of a correlation between surprise and probability of recall, and so lack of sensitivity may have been responsible for the failure of this effect to reveal itself.

Correlation between measures of curiosity. The first row in Table 3 confirms that questions marked 'K' by the experimental group in Phase I, indicating desire to know the answers, were more likely to be answered correctly in Phase III, thus revealing the predicted correlation between our two measures of curiosity. Since we find that both the recognition and the subsequent recall of answers are more probable if the corresponding questions are marked 'K', the relation between the first two may be a reflection of these correlations instead of being due to the reasons we adduced. The present experiment alone does not allow us to decide between these alternative explanations.

Surprisingness of answers. It remains to investigate the comparative ease of learning statements marked 'S' (as surprising) in Phase II and statements not marked 'S'. The last two rows of Table 4 show that the experimental group, as predicted, recalled a higher proportion of surprising statements in Phase III. But this prediction was not significantly borne out by the control group.

No legitimate conclusion can be drawn from this difference between the two groups, as the control group marked significantly more statements as surprising ($\chi^2 = 237·94$, P is less than 0·0001). This was probably due to the fact that the control group had only one task to perform in Phase II, whereas the experimental group had two. As a result, subjects in the control group are likely to have had a lower threshold of surprisingness for a statement to be marked 'S', and so their data may have been less sharply discriminating.

The effect of the surprisingness of an answer could be explained away if this variable were associated with the familiarity of the animals. But χ^2 for both experimental and control groups is under 1·00, and so this can be ruled out.

V. SUMMARY

An exploratory experiment, designed to test some of the predictions from a theory of human epistemic curiosity, is reported. An experimental group received (1) a fore-questionnaire about invertebrate animals, (2) a series of statements including answers to the questions, and (3) an after-questionnaire, repeating the questions of the fore-questionnaire. A control group underwent the same procedure except for the fore-questionnaire. The recall of the answers in the after-questionnaire and subjects' reports were used as measures of curiosity. The results tend to confirm the following predictions: (1) that pre-questioning arouses curiosity, (2) that the two measures of curiosity are correlated, (3) that statements recognized as answers to questions from the fore-questionnaire are more likely than others to be recalled in the after-questionnaire, (4) that questions about more familiar animals and questions whose concepts seem incompatible arouse more curiosity than others, (5) that surprising statements are more likely to be recalled as answers in the after-questionnaire than others.

This article is based on part of a dissertation which was submitted to the Faculty of the Graduate School of Yale University in partial fulfilment of the requirements for the degree of Doctor of Philosophy. The project was carried out as part of the Yale Communication Research Programme, which is directed by Prof. C. I. Hovland and financed by a grant from the Rockefeller Foundation.

REFERENCES

BERLYNE, D. E. (1951). Attention, perception and behaviour theory. *Psychol. Rev.* LVIII, 137–46.
BERLYNE, D. E. (1953). Some aspects of human curiosity. (Ph.D. thesis on file in Yale University Library.)
BERLYNE, D. E. (1954). A theory of human curiosity. *Brit. J. Psychol.* XLV, 180–91.
COCHRAN, W. G. (1943). Analysis of variance for percentages based on unequal numbers. *J. Amer. Stat. Assoc.* XXXVII, 287–301.
HEBB, D. O. (1951). The role of neurological ideas in psychology. *J. Personal.* XX, 39–55.
HULL, C. L. (1943). *Principles of Behavior.* New York: Appleton-Century.
HULL, C. L. (1952). *A Behavior System.* New Haven: Yale Univ. Press.
LEWIN, K. (1935). *A Dynamic Theory of Personality.* New York: McGraw Hill.
MILLER, N. E. (1948). Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. Abn. Soc. Psychol.* XLIII, 155–78.
MILLER, N. E. (1951). Comments on theoretical models—illustrated by the development of a theory of conflict behavior. *J. Personal.* XX, 82–100.
MOSES, L. E., (1952). Non-parametric statistics for psychological research. *Psychol. Bull.*, XLIX, 122–43.
NOBLE, C. E. (1952). The role of stimulus meaning (m) in serial verbal learning. *J. Exp. Psychol.* XLIII, 437–46.
NOBLE, C. E. (1953). The meaning-familiarity relationship. *Psychol. Rev.* LX, 89–98.
SKINNER, B. F. (1947). *Verbal Behavior.* (William James Lectures, Harvard University. To be published.)
SKINNER, B. F. (1953). *Science and Human Behavior.* New York: Macmillan.

(*Manuscript received* 22 *January* 1954)

BETWEEN WAKEFULNESS AND SLEEP: HYPNAGOGIC IMAGERY

By PETER McKELLAR and LORNA SIMPSON

Department of Psychology, Aberdeen University

I. *Introduction* (pp. 266–267). II. *Plan of present investigation* (pp. 267–268): (i) *Pilot-study* (pp. 267–268); (ii) *Main investigation* (p. 268). III. *Data obtained* (pp. 268–271). IV. *Discussion with further illustrative instances* (pp. 271–273). V. *Notes towards a theory of visual hypnagogic images* (pp. 273–275). VI. *Summary* (pp. 275–276). *References* (p. 276).

I. Introduction

Interesting and little-understood psychological states divide wakefulness from total sleep. These 'twilight' or 'hypnagogic' states are sometimes accompanied by imagery which in vividness and other characteristics appears different from that of both dreams and ordinary waking fantasy. In his *Dictionary of Psychology* Warren defines hypnagogic imagery as 'imagery of any sense modality, frequently of almost hallucinatory character, which is experienced in the drowsy state preceding deep sleep'.

To illustrate such imagery we shall quote the experience of one of the subjects of the investigation to be reported: 'I had images of bright green frogs which appeared to leer at me, then when they blinked their eyes, they vanished. Once I had images of evil-looking faces peeping over their shoulders at me.' (Female, aged 18 years 8 months.) Another subject's report illustrates equivalent auditory imagery: 'Sometimes I hear bells before I sleep, loud bells and tinkling bells, all ringing together.' (Female, 18 years 5 months.) Whether we call such experiences 'images', 'visions' (Leaning, 1925), or 'hallucinations' (Freud, 1900), our findings support the view stated by Leaning that the experiences are 'common', and do not accord with the low incidence found by Müller (1848), also working on a student sample. Frequent, regular hypnagogic imagery seems relatively rare, but to have experienced at least one hypnagogic image seems to be by no means uncommon. Some individuals nevertheless are unable to recall ever having had such images, and to these people a hypnagogic imager may have great difficulty in communicating his own experiences. This difficulty is discussed by Galton (1883) who, in his inquiries into visual memory, was impressed by the number of people who reported 'visions'. Galton draws attention to the intolerance which the individual with atypical, or seemingly atypical, visualizing experiences is likely to encounter at the hands of other people.

Hypnagogic images are interesting in themselves, but seem also to be relevant to a number of wider issues. Certain writers, among them Robert Louis Stevenson (Still, 1952, p. 192), and probably Edgar Allan Poe (Mowrer, 1950, p. 635n.), appear to have used their hypnagogic imagery as sources of their literary productions. The seemingly 'original' characteristics of at least some hypnagogic images may make them relevant to the understanding of certain kinds of creative thinking. Again, certain 'visions' of a kind which have been, and still are, interpreted supernaturally resemble closely what might be classified under the more neutral heading of hypnagogic images. Hobbes (1651, ch. XII), Müller (1848), and Ernest Jones (1949, pp. 58f.) have been among those who have drawn attention to the ways in which dreams and visionary experiences have contributed to the world's superstitions and folklore. The history of hypnagogic images, interpreted supernaturally as 'visions', 'omens', and 'precognitions', rather than naturalistically, is probably a long one. If Skinner's view (1953) of superstition as a product of the reinforcement of accidental contingencies is helpful, the chance occurrence of a vivid hypnagogic image with an emotionally significant event would be a potent source for superstition. It is probable that the importance of hypnagogic images in this context has been underestimated rather than overestimated. Certainly a hypnagogic image when it is experienced can mean very different things to different people. To a psychologist it may be simply an interesting phenomenon for study; to a writer or painter it may be valuable subject-matter for use; to a normal adolescent or a person who fears 'insanity' it may be a source of worry or even terror. Since hypnagogic images have resemblance to both 'visions' and hallucinations, it is relevant to record that our own interest in them arose during investigations with the drug mescaline, which is commonly used experimentally to induce psychotic-like states. While experimentally under the influence of mescaline, the first writer sought to communicate his visual experiences by searching for the closest analogy he could in experiences that the experimenters might have had. He found this analogy in hypnagogic imagery, and has since learnt that the resemblance of mescaline visions to hypnagogic images has also been independently noted by Rawcliffe (1952). Precisely because the two judgements are entirely independent, this agreement is the more interesting. If the resemblance is a true one, hypnagogic images have additional interest, since mescaline has had two main uses: (*a*) the induction of visions for mystical and religious purposes, first by the Aztecs and later by contemporary Central American Indian tribes, and, (*b*) the induction of psychotic-like states and accompanying hallucinations by experimental investigators of schizophrenia. It is interesting to consider the hypnagogic image in relation to the continuum which Galton postulates between complete hallucination and the more normal forms of imagery, particularly since, as Osmond & Smythies have recently pointed out (1952), the biochemistry of mescaline and adrenaline is very similar. Hypnagogic images merit investigation in that they may represent an instance in which greater knowledge of the 'normal' may illuminate the 'abnormal'.

It will be our task to describe instances of such images, to provide a provisional classification and tentative quantitative data, and to consider their possible relevance to the broader issues that we have indicated. We have attempted to distinguish what might legitimately be called 'hypnagogic images' from other related or superficially similar phenomena. The latter include ordinary images, dream images, hallucinations, the images that carry over from sleep into wakefulness (hypnopompic images), and the images of those to whom nitrous oxide or certain other general anaesthetics have been administered.

II. PLAN OF PRESENT INVESTIGATION

(i) *Pilot-study*[1]

The main investigation followed a questionnaire study of imagery, and related processes, conducted in 1952 on a group of Aberdeen University students, 28 male and 52 female, whose mean age was 21 years. Although the primary purpose of the questionnaire was not to elicit detailed information about hypnagogic imagery, five questions relating to this subject were included, and it was on the basis of the answers obtained that the investigators were able to design the more suitable questionnaire used in the main investigation. Despite the inadequacy of the questions, no less than 50% of the subjects reported some kind of hypnagogic imagery, the commonest being visual, though this may well have been an artefact of a visual bias in the wording of the questionnaire. Further preliminary studies were conducted on colleagues and advanced students, who were found to have an even higher incidence of hypnagogic imagery. Rejecting the improbable hypothesis that mental abnormality is unusually high among their students and colleagues, the writers approached the main investigation with the implicit assumption that they were studying a 'normal' rather than 'abnormal' psychological phenomenon.

---

[1] Reported in a paper given to the Scottish Branch of the British Psychological Society, Aberdeen, 9 May 1953. Copies of our questionnaires are lodged with the Technical Information and Documents Unit, D.S.I.R.

(ii) *Main investigation*

The aims of the investigation were to obtain descriptions of a reasonably large number of instances of hypnagogic images of different types, to look for regularities in their content and structure, and to classify them. We have been concerned more with qualitative description than with frequencies of incidence. Two groups were investigated:

Group A. Students of Aberdeen University, 110 in all, 50 male and 60 female. This group was the subject of the more 'extensive' part of the study (mean age 20 years 11 months).

Group B. A heterogeneous group of colleagues; fluent hypnagogic imagers; postgraduate students and other adults ($N = 30$). All may safely be assessed as of above average intelligence, and their introspections tended to possess both the merits and demerits of 'introspective sophistication'. This group was the subject of the more 'intensive' part of the investigations.

All figures quoted relate to group A, and not to group B: illustrative material is taken from both.

The questionnaire related primarily to hypnagogic imagery, but questions relating to other possibly relevant experiences were also asked; these included *déjà vu*, colour associations, number forms, the pre-sleep experience of 'falling', etc. It was thought justifiable to inform subjects that what was being investigated was an experience undergone by a large number of people who were not apparently maladjusted or psychotic. Despite this, it seems likely that superficial investigation by questionnaire is more likely to underestimate than to overestimate the incidence of such experiences. While reassurance may greatly minimize conscious resistance, it can do relatively little to prevent repression or 'denial' (in the psychoanalytic sense) of hypnagogic images. Information on this and on other matters of detail was furnished by group B, in interview, and by some of the subjects in group A, after they had completed the questionnaire.

III. DATA OBTAINED

Incidence. Hypnagogic imagery was reported by 67 of our 110 subjects (27 male and 40 female). For comparative purposes we may note that it was more common than colour associations (28 instances); hypnopompic images (23 instances); synaesthesia (15 instances); and number forms (8 instances). It was less common than *déjà vu* (78 instances) and the before sleep experience of 'falling and waking up with a jerk' (88 instances).

Kinds of imagery. When the hypnagogic images of the 67 subjects who had them were classified according to their sense-mode, the commonest were the auditory (50 instances), and the next most common the visual (39 instances). Kinaesthetic imagery was reported by 11 subjects, and tactile and olfactory imagery in one instance each. (One of the subjects in the pilot investigation reported a visual image with accompanying temperature imagery.) When the 67 subjects were classified according to the kinds of hypnagogic images that they had, 27 reported auditory images alone; 15 auditory and visual; 12, visual alone; 6, auditory, visual and kinaesthetic; 4, visual and kinaesthetic; 1, auditory and kinaesthetic; 1, visual and olfactory; and 1, visual, auditory and tactile. It will be noted that 39 subjects had only one kind of hypnagogic imagery, 21 subjects had two kinds, and 7 had three kinds.

Auditory imagery. Of the 50 subjects who reported auditory hypnagogic imagery, 2 described them as occurring 'regularly', 12 as occurring 'often', and 36 reported them 'occasionally'. Auditory images of music were the commonest kind, and were quite often extremely specific. Thus one subject reported: 'The most common is the last movement of Rachmaninov's Second Piano Concerto. Occasionally I hear the First Movement of his Third Concerto, which I do not know well enough to construct in my mind when I am fully conscious.' (Male, 18 years 8 months.) In 17 of the 50 instances the imagery was confused with real sounds. The following report is typical: 'I once thought I heard my mother calling me just as I was dropping off, but when I got up I found she had not.' (Female, 18 years 11 months.) This difficulty of distinguishing hypnagogic images from hallucinations was most apparent with the auditory kind. Because of the predominantly visual character of dreams the auditory images were readily distinguished from dream images.

Visual images. One of the 39 subjects who reported visual hypnagogic images had the experience 'regularly', 4 had the experience 'often', and 34 'occasionally'. The images were coloured in 18 instances, uncoloured in 14, and unstated in 7 cases. The vividness and other characteristics of the colour experiences seemed to impress greatly a number of the subjects. Visual hypnagogic images were ordinarily experienced with closed eyes, but open-eyed images have been reported (e.g. Rawcliffe, 1952). Our data confirm this, since in 6 cases subjects experienced them with open eyes. The open-eyed hypnagogic image of 'the room being full of angels' reported by one subject, particularly as it immediately preceded the death of her husband, is the type of experience likely to be given a supernatural interpretation, as it was in this instance. Other instances, however—for example, the open-eyed image of 'a ship in a storm—the mast was broken and the sails were torn—tossed by huge waves'—did not involve the kind of 'accidental contingency', which, following Skinner (1953), we have suggested would probably play an important part in supernatural interpretations of natural phenomena. It is very difficult to draw any sharp line between such open-eyed 'hypnagogic' experiences, 'eidetic' images, and actual 'hallucinations'. Whatever we call them these visualizations represent a sort of analogue to the auditory images which were confused with real perceptions.

The content of the visual hypnagogic images was extremely varied and such as to evoke a variety of emotional reactions. To one subject his images were 'a background to sleep, almost in fact a lullaby'; to another they were a source of amusement and pleasure, their content being rather like that of a Walt Disney cartoon; to others they were a source of acute distress and even terror, being, as one subject reported, 'a terrible jungle through which I have to pass before sleep comes'.

Differences from dreams. The subjects were asked to say in what ways their hypnagogic images differed from their dreams. Typical answers were that the images were more vivid and realistic, that they 'come and go in a flash', and that the subject did not participate in them as in a dream, but 'only watched'. One subject described his images as 'like a succession of vivid lantern slides appearing without voluntary control, and containing detailed material which I didn't know I knew'. (Male, 34 years 7 months.)

Subjects were also asked how they knew that they were awake and not asleep when experiencing the images. Since the question is a difficult one, several were unable to give a satisfactory answer, and there is always the possibility that hypnagogic images, particularly those of a visual kind, may be confused with those of dreams. Among the reasons given for believing oneself to be awake were: being able to have ordinary perceptions at the same time (the commonest reason given); being able to have separate thoughts; being able to engage in conversation; being able to open eyes, close them, and continue with the image, etc. The following instance is far from atypical in the definiteness with which the subject was aware of being awake: 'It was dark outside, but I "saw" out of the window—bright sky, coloured trees. I could see people sitting on the roof—what they were wearing and the features of their faces. I woke my sister who was sleeping beside me and said "Look, Ella!" and began to count the people.' (Female, 19 years 3 months.)

Differences from ordinary images. To the question in what way their hypnagogic images differed from ordinary images, two subjects reported no difference, and several were unable to respond. The majority, however, were able to put the difference into words. They described their hypnagogic images as being clearer and more vivid; less subject to conscious control; brighter in colour; 'more a train of events than a single image which can be called up in the mind, and one thing leads to the next'; 'I could feel I wasn't exactly doing the things'; and other reasons. Gordon (1949, 1950) has developed the idea, stemming from the work of Jaensch (1930) on eidetic images, that ordinary waking images can be divided into those which are 'autonomous', and those which are of a voluntary controlled kind. Hypnagogic images seem to be largely, if not exclusively, of the autonomous kind. At an early date Müller (1848) stressed the autonomy of his own hypnagogic images and the fact that he had no control over either their production or their changing form. The visual images closely resemble dreams in this characteristic, which, as Gordon's work shows, is also possessed by some but not all ordinary images.

Definition of hypnagogic images. Subjects were given this definition of hypnagogic images: 'Imagery of any sense, frequently of intense, almost hallucinatory vividness, experienced in the drowsy state before sleep.' They were asked to state whether this adequately described their own experiences, and if not in what way their images differed. Of the subjects, 33 agreed it was a good definition, 21 disagreed in some particular and 13 had no comment. The reason for the disagreement most commonly given was that the image was not as vivid as the definition implied (18 instances): 'never intense or hallucinatory', and 'not particularly vivid, but always extremely persistent' are representative comments. Other subjects supported the interesting observation made by Collard (1953) that drowsiness and subsequent sleep were not invariably necessary to the occurrence of the imagery. One subject reported: 'I can be in bed wide awake or even have them in the ordinary course of the day if alone.' There are probably wide individual differences in this matter, and our data yield definite evidence that in some cases drowsiness is not a necessary component. Perhaps, however, these deviant cases are better described as 'eidetic' than as 'hypnagogic' images.

Conditions favourable to hypnagogic images. It will be noted that some subjects experienced hypnagogic images regularly and without any special precipitating factors. We should not wish to set much store by the subjects' statements about the causes of their images. As might be expected, words like 'worry' and 'fatigue' were frequently used, but these are highly ambiguous and likely to be merely rationalizations and stereotyped ways of responding to a question on a matter of this kind. Definite evidence regarding the mental conditions that favour hypnagogic imagery must await further investigation, which may disclose wide individual differences.

IV. DISCUSSION WITH FURTHER ILLUSTRATIVE INSTANCES

Readers interested in the history of investigations of hypnagogic imagery are referred to the scholarly study by Leaning (1925). Leaning, who draws attention to the lack of investigations by psychologists in this field, surveys alternative explanations and lends his own support to a supernatural rather than a naturalistic theory. Any alternative naturalistic explanation, such as we should look for, must do justice to the fact that imagers often find the experiences extremely impressive in their apparent originality—extremely strange and foreign to the imager's personality. We should certainly wish to endorse Leaning's contention that the imager's own beliefs greatly affect his interpretation of his experiences. The following hypnagogic image lends itself to supernatural interpretation, granted appropriate beliefs. 'I see a figure draped in white approaching my bed. It is vague but it is the figure of a woman and I get the impression she is a nun.... The figure I see sometimes speaks, or rather seems to be saying my name.' (Male, 18 years 10 months.) Two of the subjects in group B were a mother and a daughter, both strongly hypnagogic. The mother interpreted her images supernaturally, the daughter in naturalistic terms; both were very decided in their opposing interpretations.

What is true about the influence of antecedent beliefs on the interpretation of hypnagogic images also applies to psychotic beliefs. A psychotic might readily interpret a hypnagogic image in terms of his delusions, and if, as seems highly probable, hypnagogic images have enriched the folklore of mankind, they might similarly enrich the content of a pre-existing psychosis. A curious paradox arises, that to understand abnormal mental occurrences more fully, it may be necessary to examine abnormalities themselves in terms of *seemingly* abnormal experiences which a psychotic merely shares with many non-psychotic normal individuals. Consider the image reported by a student: 'I see skeleton figures, firstly very thin, then very fat. They appear alternately. The thin ones usually appear white and the fat ones black. I see this very often.' (Female, 20 years 2 months.) This is only one of the numerous frightening hypnagogic images reported, and a psychotic might also be frightened by such a psychotic-like experience, even although it is shared by normal people.

The fact that hypnagogic images can provide an individual with experiences greatly different from those he might otherwise have may be noted. One category of hypnagogic images involved the experience of change of size of the subject's own body or of things seen. One subject reported seeing people known to him who 'always grow bigger and bigger until I cannot picture them, or fade away to a single point'. (Female, 20 years

2 months.) Another subject reports a kinaesthetic image: 'all my body was swelling like a balloon and was being held together by thin bands of string in place of bones'. (Male, 22 years 11 months.) He states that the experience often preceded a standard nightmare. These images permit an individual to undergo experiences like those in a fanciful children's story—for example, two well-known incidents in *Alice in Wonderland*. They are also similar to the sensory experiences that form the basis of psychotic delusions of bodily change. Similar experiences occur to investigators who have experimentally induced psychotic-like states in themselves with such substances as mescaline and lysergic acid. The first writer after being given 0·5 gram of mescaline experienced a distinct sensation of elongation of one side of his body, this incidentally being reflected in the human figure drawings he produced at the time. Again, one subject who had taken an unstated dose of lysergic acid reported: 'As I watched, my fingers grew into my hand, my hand telescoped into my arm...my legs were now shrinking...my head, which was now double its normal size....' (Katz, 1953.) The possibility that a size-distorting mechanism, not unlike that occurring in psychosis and in experimentally induced psychotic-like states, operates for some subjects during the process of falling asleep merits further investigation.

It is possible to divide visual hypnagogic images into two classes. For want of better terms we shall refer to these respectively as 'perseverative' and 'impersonal'. The perseverative images are those easily located in past experience. A typical instance is the hypnagogic image of 'sometimes a clear picture of someone I had perhaps met during the day but I am not thinking about at the time'. (Female, 18 years 9 months.) The impersonal images differ; they are not easily explained in terms of past experiences, and may seem interestingly or alarmingly foreign to the personality. One subject reports 'a large bloated yellow head, pouting red lips, wild pale blue eyes rolling, hair dishevelled.... The mouth was moving but there was no sound'. (Male, 21 years 8 months.) Another subject reported a similarly unexplained and similarly frightening series of faces experienced with open eyes. Some were impressed with the images as artistic subject-matter. Thus one subject reports visual images 'of a scene which I feel I would like to paint—these scenes are usually original'. (Female, 22 years 11 months.) These impersonal images, at times almost psychotic-like in their content, present difficulties of explanation that do not arise with the perseverative type of image. It is, nevertheless, our contention that these hypnagogic images, like other psychological occurrences, are subject to the principle stated by Hobbes: 'There is no conception in a man's mind which hath not at first, totally or by parts, been begotten upon the organs of sense.' (*Leviathan*, pt. 1, ch. 1.) With some success in a number of instances we have sought to defend this principle against 'archetypal', supernatural or other unsatisfactory hypotheses, by studying impersonal images with free association. An illustration may be taken. The subject, who was himself a clinical psychologist (age 31 years 3 months), reported in interview the following hypnagogic image: 'A beautifully formed and perhaps varnished claw, and only a claw. I suspect there was a foot originally and that it turned to a claw. Vividly and accurately coloured—a much more accurate claw than I could draw, unless I could use oils. Background appropriate but couldn't describe. The claw was outstanding. A cloud effect the base of the picture—like backs of sheep—nothing definite—focus of attention on the claw.' This description was taken down verbatim without interruption, and the subject was unable to locate the content of the image in past experience. A period of free association with occasional questions followed. Space prevents giving the associations in detail but they included; 'the claw has dog's hair...the dog lying on the mat in front of fireplace with extended claws...base of picture could have been bank of cloud seen from aircraft'. The subject reported having done quite a lot of flying recently. He was now able to recall that a few days before he had seen a stuffed animal's head with highly polished horns. The mounting of the claw was similar to the mounting of this head. He also recalled a surrealist painting with something in it corresponding to the position of the claw in the image. Eventually the subject related a personal problem which had been causing considerable irritation. With his wife and two children, he had been staying with a relative; he had been repeatedly resentful of the sentimental attitude of the relative to her dog, and the preferential treatment the dog had received in any conflict between it and the children. The inquiry by free association eventually resulted in the reduction of the hypnagogic image to a dramatization of the threat to the children which the dog represented. The dramatization was built up out of past experiences. While this instance, like others also studied by free association, does not prove the hypothesis that images derive ultimately from sense perception, it at least illustrates the possibility of its validity. It also indicates that when conscious recall is not able to provide explanation of either the motivation or the content of a hypnagogic image, subsequent free association may help to supply this omission. The possibility that the hypnagogic image is like the dream in being a valuable clue to the personality and its emotional preoccupations may prove an extremely fruitful point for future investigation.

V. NOTES TOWARDS A THEORY OF VISUAL HYPNAGOGIC IMAGES

It has been suggested that the content of the hypnagogic image is provided by perceptual experiences. It is further suggested that the explanation of hypnagogic images must take account of the sense organs and central nervous system of the experiencing subject. It is now established that the brain has a spontaneous electrical activity, and that sense impressions are occurrences *superimposed* upon this spontaneous, autonomous activity which goes on all the time. As well as the spontaneous activity of the brain there is the similar activity of the sense organ, in this instance the retina of the eye. Without any external stimulation this spontaneous activity may yield a changing sequence of shapes and hues. As William James puts it: 'The retina itself, even in complete objective darkness, seems to be always the seat of internal changes which give some luminous sensation.' (1892, p. 46.) In a most interesting passage Galton (1883) candidly admits his own inability for a very long time to recognize this fact: 'I should have emphatically declared that my field of view in the dark was essentially a uniform black....I have found this is by no means the case, but that *a kaleidoscopic change of patterns and forms is continually going on*, but they are too fugitive and elaborate for me to draw on with any approach to truth. I am astonished at their variety.' (P. 114, italics ours.) A very careful study was made by Ladd (1892) of these products of retinal activity to which Wundt had given the name of 'luminous dust'. (They are alternately called '*Eigenlicht*', 'idio-retinal sensations', and 'phosphene' experiences.) Ladd reported: 'By far the purest, most brilliant and most beautiful colours I have ever seen, and the most astonishing combinations of such colours, have appeared with closed eyes in a dark room.' (P. 300.) It was after careful introspective study of these spontaneous retinal processes in an attempt to verify

Müller's theory of their relation to sleep that Ladd came to accept the view that the forms and colours constituting 'luminous dust' provide the essential basis of dream images. The question arises whether some extension of this theory might not be applicable also to hypnagogic images of the visual kind. The unusual and unexpected forms which hypnagogic images can assume present a difficulty for any explanation wholly in terms of reorganizations of *past* perceptual experiences. Some explanation more compatible with established psychological and biological knowledge than eruptions of a 'collective unconscious' and its attendant 'archetypes' seems desirable. This alternative is perhaps to be found in the interaction of present sensory experiences of the 'luminous dust' kind with past perceptions and their reactivation. Of such luminous dust Ladd reports: 'I verily believe *there is no shape known to me by perception or by fantasy*, whether of things on earth or above the earth or in the waters, *that has not been schematically represented by the changing retinal images* under the influence of intra-organic stimulation.' (P. 300, italics ours.)

Four characteristics of the visual hypnagogic images we have encountered in our investigation seem to require special explanation. There is, first, their frequent reported vividness; secondly, their autonomous character and independence of conscious control; thirdly, the original (unreproductive) character which at least some of them possess; and, fourthly, their ever-moving, or ever-changing, nature. If contemporary sensory processes play a part in the production of hypnagogic images (and, perhaps, as Müller and Ladd have suggested, of dream images), this would help to explain their vividness. The immense variety and rapid change in luminous dust to which Ladd refers—occurrences that are wholly outside voluntary control—would help to explain the autonomous and seemingly 'archetypal' character of visual hypnagogic imagery. It is not necessary to rest any such explanation wholly upon spontaneous retinal processes as opposed to those of a stimulus-induced kind. Inadequate stimuli such as pressure, as can be readily verified, will greatly augment these retinal sensations. As Müller pointed out, when presenting an earlier version of the theory of dream images which Ladd developed, during sleep 'both eyes are turned inwards and upwards' (1848, p. 1413). The adjustment of the eyes in going to sleep would be likely to produce a considerable amount of inadequate stimulation of a kinaesthetic kind. Nor can we exclude the possibility that other physiological changes, attendant on the process of falling asleep, may also provide inadequate stimuli, though of an internal kind, either to the retina or to the visual areas of the brain. Müller points out (1848, p. 1343) that inflammation of the optic nerve will produce 'luminous phantasms'. Presumably, biochemical stimulation gives rise to the visions produced by such drugs as mescaline and lysergic acid. The reader is reminded of the similarity noted between mescaline-induced visions and the first writer's own hypnagogic images.

The kind of relation envisaged between luminous dust (with these possible augmenting sources of inadequate stimulation) and hypnagogic images may be indicated. Hypnagogic imagery, we suggest, does not occur against a screen of darkness, but against a changing screen of patterns and colours on to which the central processes are 'projected' (the term is used in an analogous sense); and this 'screen', being a non-homogeneous field of changing patterns, contributes to the resulting image. The outline of the images is provided by the luminous dust, but these are interpreted on the basis of such past perceptions as are reactivated. The situation indicated is more like highly personal interpretations being given to unstructured sense data (as in a projective test) than to simple free associative activity in the absence of external determinants. Given such an interaction of central processes (past perceptions) with sensory processes of the retinal-dust kind (present perceptions), it seems possible to explain both the unexpected and impersonal character of such imagery, with the possibility that it may nevertheless provide information about the subject's personality, as do dreams. It is widely accepted that both exteroceptive and interoceptive sensations, as well as associative processes, play some part in the dream. That contemporary sensory processes, namely retinal dust, contribute substantially to the explanation of visual hypnagogic imagery seems a reasonable hypothesis for further investigation.

Visual hypnagogic images are sometimes, though rarely, found to occur with open eyes. Luminous dust may also be relevant to such open-eyed experiences. It has been noted that Galton postulated a continuum from ordinary imaging to hallucination. A little investigation showed the writers that retinal dust, continuous with that experienced with closed eyes, can be detected in twilight conditions with open eyes. The fact that apparitions are perceived or misperceived more readily in twilight than in broad daylight may be noted. It is possible to proceed further, and careful introspection reveals that luminous dust of an achromatic kind—a sort of mottled atmospheric haze—can be perceived even in broad daylight. This is difficult to observe, since the perceptual equivalents of spontaneous retinal activity are ordinarily suppressed by daylight. In Gestalt terms, they form an extremely insignificant 'ground' to the strong 'figure' which perceived objects present. Retinal dust is so omnipresent as to be readily overlooked, and was overlooked (even with closed eyes) for a long time by as acute an introspectionist as Galton. William James formulates a principle that may help to summarize our argument: 'Any objective stimulus to be perceived must be strong enough to give a sensible increment of sensation over and above the idio-retinal light.' (1892, p. 46.)

To this principle we would add that, when such increment is wholly or mostly absent (in darkness or with closed eyes), idio-retinal light is conspicuous, while in daylight it is barely discernible. We have suggested the probable relevance of such ideo-retinal light to hypnagogic images, and Ladd has suggested its relevance to dream images. The possibility remains that, even in the presence of stimuli that give sensible increment, it may exert an effect. What is ordinarily ground may become figure, and in semi-darkness may give rise to apparitions and other such experiences interpreted supernaturally. There remains the possibility of its relevance to hallucinations in full daylight. Cameron (1947, p. 542) reports the tendency of a delirious patient to experience 'dream-like hypnagogic hallucination when he shuts his eyes' before passing to actual hallucination with open eyes. Galton suggests that some very slight physical change 'short of really morbid character' would enhance the vividness of luminous dust experiences. This suggestion is interesting, since in the typical kinds of psychosis in which visual hallucination is a prominent symptom there are known to be biochemical or other physiological processes at work; these might well supply inadequate stimulation sufficient to induce the postulated figure-ground reversal.

VI. SUMMARY

An investigation of hypnagogic images was conducted on three groups of Aberdeen University students and staff. Provisional quantitative data are given, but the main concern of the investigation has been to collect descriptive accounts of hypnagogic images, to classify these images, and to outline a theoretical explanation of those of a visual kind. It is suggested that hypnagogic imagery may throw light upon a number of wider issues including some forms of creative thinking, and various paranormal phenomena, notably 'visions'. Data suggest that hypnagogic experiences may help to fill in gaps in the continuum which Galton supposed to extend from complete inability to image to gross hallucination. Hypnagogic images of an auditory type—the commonest kind in the group studied—are not readily distinguishable from hallucinations; visual hypnagogic images both resemble and differ from dreaming and ordinary waking images. The theory that the spontaneous activity of the retina plays an important part in visual hypnagogic imagery—a development of Ladd's theory of the dream—is considered and defended as helping to account for a number of characteristic features of the images.

REFERENCES

CAMERON, N. (1947). *The Psychology of the Behaviour Disorders.* New York: Houghton Mifflin.

COLLARD, P. (1953). Hypnagogic visions. *Light*, LXIII, no. 3401, 233–5.

FREUD, S. (1900). *The Interpretation of Dreams.* London: Allen and Unwin.

GALTON, F. (1883). *Inquiries into Human Faculty and its Development.* London: Everyman.

GORDON, R. (1949). An investigation of the factors that favour the formation of stereotyped images. *Brit. J. Psychol.* XXXIX, no. 3, 156–67.

GORDON, R. (1950). An experiment correlating the nature of imagery with performance on a test of reversal of perspective. *Brit. J. Psychol.* XL, nos. 1 and 2, 63–7.

HOBBES, T. (1651). *Leviathan.* Edited Molesworth, W., 1839. London: Bohn.

JAENSCH, E. R. (1930). *Eidetic Imagery.* London: Kegan Paul.

JAMES, W. (1892). *Psychology: Briefer Course.* London: Macmillan.

JONES, E. (1949). *On the Nightmare.* London: Hogarth.

KATZ, S. (1953), *Macleans Magazine.* October 1, 1953.

LADD, G. T. (1892). Contribution to the psychology of visual dreams. *Mind* (New Series), I, 299–304.

LEANING, F. E. (1925). An introductory study of hypnagogic phenomena. *Proc. Soc. Psych. Res., Lond.*, XXXV, 289–409.

MOWRER, O. H. (1950). *Learning Theory and Personality Dynamics.* New York: McGraw Hill.

MÜLLER, J. (1848). *The Physiology of the Senses.* London: Taylor, Walton and Maberly.

OSMOND, H. & SMYTHIES, J. (1952). Schizophrenia: a new approach. *J. Ment. Sci.* XCVIII, 309–15.

RAWCLIFFE, D. H. (1952). *The Psychology of the Occult.* London: Ridgway.

SKINNER, B. F. (1953). *Science and Human Behaviour.* New York: Macmillan.

STILL, A. (1952). *Borderlands of Science.* London: Rider.

WARREN, H. C. (1934). *Dictionary of Psychology.* New York: Houghton Mifflin.

Note. An historical treatment of hypnagogic imagery from the time of Aristotle is to be found in Leaning (1925), who also gives a comprehensive bibliography.

(*Manuscript received* 4 *January* 1954)

SOME EFFECTS OF INCENTIVES ON THE PERFORMANCE OF IMBECILES[1]

By S. GORDON, N. O'CONNOR and J. TIZARD

Medical Research Council, Unit for Research in Occupational Adaptation

I. *Introduction* (p. 277). II. *Some assumptions* (pp. 277–278). III. *Selection and description of the sample* (pp. 278–279). IV. *Leg persistence: experiment one* (pp. 279–285): (i) *Experimental design* (pp. 279–280); (ii) *Experimental conditions* (p. 280); (iii) *Matching of groups* (pp. 280–281); (iv) *Results* (pp. 281–282); (v) *Conclusions* (p. 282). V. *Leg persistence: experiment two* (pp. 282–285): (i) *Introduction* (pp. 282–283); (ii) *Experimental design* (p. 283); (iii) *Results* (pp. 283–285); (iv) *Conclusions* (p. 285). VI. *Leg persistence: experiment three* (pp. 285–286): (i) *Introduction* (p. 285); (ii) *Method* (p. 285); (iii) *Results* (pp. 285–286); (iv) *Discussion* (p. 286). VII. *Conclusions* (pp. 287–288). *References* (p. 288).

I. INTRODUCTION

Although a few investigators have carried out experiments with imbeciles in which incentives have been used (Aldrich & Doll, 1931; Gardner, 1945) the problem of how low-grade defectives respond to incentives remains essentially unexplored. It has been suggested that imbeciles constitute a separate clinical entity and that certain basic operants of behaviour found among normals or even feeble-minded persons do not apply to them. Among others, Lewis (1948), and Kallman (1951), maintain that while high-grade defectives form the lower end of the normal distribution curve of human intelligence, low-grade defectives do not fit into any scheme of continuous variation. Whether qualitative differences in intelligence affect all aspects of behaviour is, however, not known.

With the general aim of contributing towards a more systematic psychology of the imbecile defective this investigation seeks to describe how imbeciles[2] respond to different incentives, and to determine to what extent their behaviour in a test situation is modifiable by experience or incentive.

In order not to have to cope with the difficult problem of complex variables involved in previous training, it was decided to select a task which was essentially unrelated to prior experience and which could be measured precisely. A measure of persistence, operationally defined as the length of time an individual maintains an uncomfortable or tiresome position, seemed suited to our purpose and the Leg Persistence test was chosen to meet these requirements. The test requires the subject to sit on a chair and hold one of his legs horizontally above another chair, keeping the heel of his shoe about an inch from the seat of the second chair for as long as he can. The choice of the Leg Persistence test was only incidentally related to investigations on persistence as such.

II. SOME ASSUMPTIONS

We assumed that the Leg Persistence test was of no intrinsic interest to the subject; that is, the subject would not enjoy doing the task itself. He would probably be willing to comply with instructions either to please the examiner or, perhaps, to avoid getting into trouble if he believed that the examiner was a person in authority. Over and above motivations such as these, it was expected that imbeciles—if they responded to incentives in a manner resembling normals (as described for example by Mace, 1935)—would modify their performance of this task according to the following pattern. Those working towards an adjusted and ascertainable goal would perform significantly better (i.e. persist longer) than those who were constantly encouraged to do their best; and both would be significantly superior to a control group. However, it was problematical whether imbeciles would be able to comprehend the significance of a goal as such and it was thought therefore that they might respond better to encouragement than to knowledge of results or self-competition.

---

[1] The present study is part of research in social and psychological aspects of mental deficiency carried out under the direction of Prof. Aubrey Lewis. Some of the work reported in this paper was presented for the Ph.D. degree at London University by S. Gordon.

[2] The sample used in the investigation reported here includes only persons about whom there is little doubt regarding their classification as imbeciles, whether one uses the criteria employed by Lewis (1929), Doll (1946), or Tredgold (1947).

III. SELECTION AND DESCRIPTION OF THE SAMPLE

The subjects were twenty-seven males certified as imbeciles under the English Mental Deficiency Laws and resident patients in a large hospital for mental defectives.

They were selected in the following manner. The case histories of patients in the two wards of the hospital which contained the largest number of male imbeciles were examined —some sixty in all. All those not doubly handicapped, such as the blind, severely crippled or epileptic, and who were between 17 and 35 years of age were selected to participate in the experiment. No attempt was made to eliminate those said to be emotionally unstable or psychotic. The majority of those not included in the experiment were patients above the age of 35. In all, thirty-one patients were individually given the Leg Persistence test. Four patients (of five diagnosed as psychotic) either did not understand the instructions or refused to co-operate and were not included in the experiment.

The age,[1] data on the number of years spent by these patients in institutions for mental defectives, and their Terman Merrill I.Q.'s, are summarized in Table 1. The subdivision into three groups is explained in section IV (p. 279). The mean age was 26·89 years, S.D. = 3·74 years. All but a few had spent the greater part of their lives in an institutional environment. The 1937 Stanford Revision of the Binet test (Form L) was given to each

Table 1. *Subjects who took part in Leg Persistence experiments*

Means, range, and standard deviations for Age, number of years spent in institutions and Terman-Merrill I.Q. of each group of subjects, and for all subjects.

Group	Mean	Range	S.D.
		Age	
Control	28·78	21–35	4·30
Encouragement	26·11	23–32	2·86
Goal	25·78	21–31	3·14
Total	26·89	21–35	3·74
		Years in institution	
Control	16·67	10–22	3·83
Encouragement	18·89	7–24	4·86
Goal	19·0	14–24	3·78
Total	18·19	7–24	4·38
		I.Q.	
Control	33·33	25–42	6·61
Encouragement	34·88	26–41	5·06
Goal	37·66	31–48	4·88
Total	35·29	25–48	5·86

---

[1] Given to the nearest whole year as at the beginning of this experiment in October 1952.

subject after the entire investigation had been completed. The results were in keeping with the diagnosis of imbecility in all cases. The mean I.Q. was found to be 35·29, S.D. = 5·86, with a range of 25–48.

Most of the subjects were described in their case notes as being 'capable of only limited work under constant supervision'. Apart from the classification of imbecility, only a few of them appeared to be of well defined clinical types. Two were mongols, and four or five were almost certainly brain injured. Three were believed to have had at least one defective sibling but as far as could be determined none had parents who were defectives. One patient was described as schizophrenic and four others as emotionally unstable. The case notes indicated that developmental retardation was apparent in all cases and had been noted by the parents or guardians before the fourth year and in most cases by the third year.

None of the subjects was able to answer correctly the questions, 'How many do 3 and 4 make?' and, 'If you had three sweets and four sweets, how many would you have altogether?'

IV. LEG PERSISTENCE: EXPERIMENT ONE

(i) *Experimental design*

Each subject was individually given the Leg Persistence test. The procedure was demonstrated by the examiner before the first trial. The subject was seated with a chair opposite him and the following instructions were given: 'When I give you the signal I want you to hold up your leg above this chair as long as you can without touching the chair. Remember, hold it up as long as you can without touching. Ready? Go!' No further comment was made until the leg touched the chair. The examiner then said, 'Good work, I'll see you again tomorrow'.

The score recorded was the number of seconds the subject held his leg in the air. A second trial with the same instructions was given the following day. Although the subjects were not given any instructions as to which leg to use, it was noted that all subjects used the same leg each time.

The higher score of the two trials was employed as a basis for dividing the twenty-seven subjects into three matched groups of nine persons each. (No other variable was used in the matching.) The subjects were ranked from best to worst according to their score on the test. The three subjects with the highest scores were grouped and each one was assigned at random to one of the experimental conditions given below. Similarly the following three were grouped and the same procedure of assigning each one at random to one of the experimental conditions was repeated, until the twenty-seven subjects were assigned to one of three experimental groups.

The Leg Persistence test was given individually as before to members of the control and two experimental groups once a day for ten consecutive days, excepting a week-end break after the fifth day. The groups were differentiated in the following manner:

Control

The original instructions. No change.

Encouragement.

In addition to the original instruction some remark of encouragement was made by the examiner every 5 sec. By the third trial this followed a quite stereotyped pattern;

'Keep it up.... Keep it up as long as you can.... That's good.... You are doing fine.... You want to try to keep it up as long as you can....'

Goal (sometimes referred to as knowledge of results).

The subject was shown a piece of cardboard (8 × 20 in.) on which was fixed a movable pointer. This could be manipulated by the examiner. A scale running from zero to 300, marked on the side of the cardboard, like the markings on the side of a mercury barometer, was used to indicate the number of seconds that elapsed. For the first experimental trial a paper clip was placed at a point 10% higher than the subject's best previous performance (original matching score). On the second and subsequent trials the clip was placed at a point 10 sec. above their best previous performance. The examiner pointed to the clip and said: 'This is as far as you got the last time. See if you can hold your leg up until this pointer passes the clip. Let us see if you can beat your record. Ready? Go!' The pointer was moved up from the zero point by a string pulley at 5 sec. intervals according to the marked gradations. The subject was told that no talking was permitted during the experimental period and the examiner himself made no further comments. At the end of the session if the subject reached his goal he was told, 'You reached your goal today. Good show. We will see if you can reach it again tomorrow'; and if he did not reach the goal, 'You did not reach your goal today. You will try again tomorrow'. On the third trial and for each subsequent trial the procedure was again slightly modified. In order to avoid frequent failures, success was contrived if the subject was within 25 sec. of his goal. This was done by pulling the pointer up to the clip just as the subject seemed ready to put his leg down.

(ii) *Experimental conditions*

A private room adjoining the wards was available for the experiment. Each subject was tested shortly after breakfast or after dinner. There were no obvious distracting influences during the experimental period. Discussion of any kind during the experiment was discouraged and all questions, of which there were many, were answered evasively. To a variety of questions such as, 'What is this for?' the reply was, 'You will see later on'.

The examiner was very well received by the patients. Rapport was good in all cases. The Charge Nurse reported that most of the subjects looked forward to the coming of the examiner each day.

(iii) *Matching of groups*

The data, as summarized for each group, for the unmatched variables of age, years spent in institutions, and I.Q. are presented in Table 1. The three groups are well matched on these uncontrolled variables, and none of the differences between groups were significant.

An analysis of variance[1] was carried out to determine whether there were any significant differences between the three groups in their initial scores on the test (the controlled variable). The means and standard deviations were respectively (Goal, $\bar{X} = 92{\cdot}56$; S.D. $= 27{\cdot}42$; Encouragement, $\bar{X} = 87{\cdot}89$ S.D. $40{\cdot}40$; Control $\bar{X} = 89{\cdot}89$ S.D. $= 41{\cdot}34$.

For 2 and 16 degrees of freedom the F ratio was 0·306 and the null hypothesis was therefore accepted. There were very significant within-group differences, the mean score for the highest performer in each group being 163 sec., while the mean for the lowest performer was 37 sec. There was a low but positive correlation ($r = 0{\cdot}355$ between initial scores on the Leg Persistence test and I.Q.

---

[1] All analyses of variance are for matched groups and follow the procedure given in Edwards (1950), pp. 274–6.

(iv) *Results*

Trials 1–10

The introduction of incentives effected striking differences between the three groups which began at virtually the same level. (The results, as indicated by the actual mean scores for each trial are presented graphically in Fig. 1, according to groups.) The Control group averaged from 85 to 101 sec. daily, the Encouragement group from 89 to 134 sec., and the Goal group from 146 to 257 sec. An analysis of variance computed for the mean score of each subject (trials 1–10) in the original matching order, as well as for the mean scores of each group for each of the ten trials, showed that the differences between the groups were very significant at less than the 0·001 level of confidence when the mean score of each subject was used ($F = 9{\cdot}895$ for 2 and 16 D.F.). (Between Goal and Encouragement, $t = 5{\cdot}465$, significant at the 0·001 level and between Encouragement and Control, $t = 2{\cdot}415$, significant at 0·05 level of confidence.)

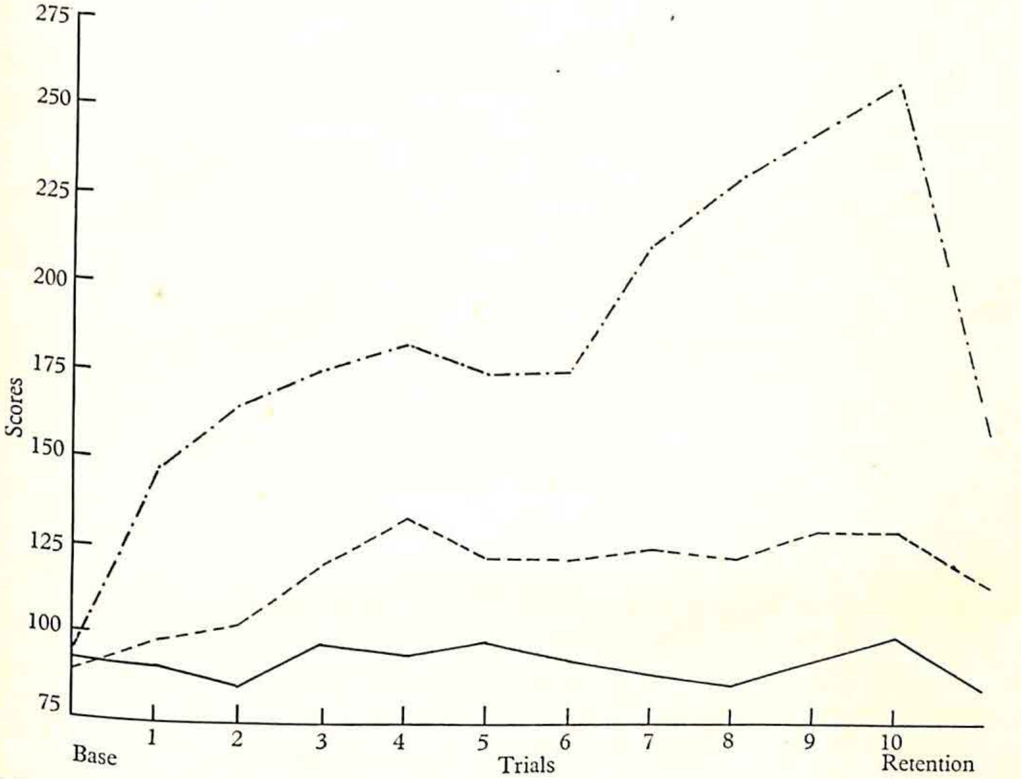


Fig. 1. *Leg Persistence*: mean score base, trials 1–10, and retention for Control, Encouragement and Goal groups. ——, Control; – – – –, Encouragement; –·–·–·, Goal.

As can be seen in Fig. 1, the Control group made virtually no progress throughout and could almost be represented by a horizontal line. After moderate gains for the first four trials the Encouragement group declined slightly and then maintained a fairly constant level of at least 25 sec. above Control for each trial. The Goal group showed on the first trial a very sharp acceleration to a point which exceeded the highest score obtained by its nearest rival—the Encouragement group. On the tenth trial the Goal group obtained a mean score which represented a gain of 123 sec. or 91% over the highest mean score obtained by the Encouragement group; and a gain of 156 sec. or 153% over the highest mean score in the Control group.

Other measures illustrate further the pattern of results obtained. The mean percentage of improvement represented by the average score for the ten trials compared with the original matching scores was 2·22 for the Control, 46·44 for the Encouragement, and 115·0 for the Goal group. When the highest single score obtained by each subject was used to compute the percentage of improvement, the mean percentage of improvement for each group was as follows: Control 35·10, Encouragement 85·22, and Goal 190·33.

No subject in the Goal group failed to improve less than 54% (mean score trials 1–10 compared with matching scores) and the greatest improvement was 209% achieved by a subject who was originally in the last matching. In the Control group the highest percentage improvement was 24% and three subjects obtained mean scores for trials 1–10 which were lower than their original matching scores. The range of percentage of improvement for the Encouragement was 0–98. The percentage improvement appeared largely unrelated to the tested intelligence. The correlation between I.Q. and per cent improvement was $r=0{\cdot}240$ in the Control group, 0·099 in the Encouragement group, and 0·138 in the Goal group.

The rank order of subjects in the Control group remained almost the same throughout the experimental period. The correlation between the original matching score and the mean score for the first five trials was $r=0{\cdot}97$ and for the second five trials $r=0{\cdot}97$. The matching scores thus appeared to be highly reliable. When incentive conditions were changed the original rankings were also somewhat changed. The correlation between the base and the mean scores for the first and second weeks in that order were as follows: Encouragement $r=0{\cdot}87$ and 0·80; Goal $r=0{\cdot}79$ and 0·82.

(v) *Conclusions*

The results suggest that imbecile subjects, engaged in a tedious task of short duration, are capable of responding to incentives in an orderly and predictable way. Those striving to reach a goal related to their standards of achievement will perform consistently and significantly better than those who are simply encouraged, however persistently, to do their best, and both of these groups will show superior results to a control group. In their manner of response, imbeciles, it would seem, do not differ from normals (as suggested by some of the experiments of Mace). This conclusion must be tempered by the consideration that although the matching on the basis of performance appeared to be sound (as substantiated statistically) there is always a possibility, especially in a small sample, that some unknown variables have been operating to affect performance scores. Consequently it seemed desirable to carry out further experiments designed partly to check the original findings and partly to elaborate and refine our conclusions.

V. LEG PERSISTENCE: EXPERIMENT TWO

(i) *Introduction*

An experiment in which some members of each group were shifted to different conditions seemed a feasible way of examining more critically our previous results. If, for example, subjects from the Control group who changed over to Goal conditions improved their performance, it would lead us to reject the possibility that the original results could be accounted for by errors of sampling. The experiment would also provide a basis for examining imbecile behaviour as affected by new incentive conditions.

(ii) *Experimental design*

The subjects, in each group separately, were ranked according to their average performance in the ten trials of the first experiment. Persons ranking 1, 3, 5, 7, 9 in the Goal and Control groups were selected to remain under the same conditions. The remaining subjects (ranks 2, 4, 6, 8) in the Control group were shifted to Goal conditions and those remaining in Goal were shifted to Encouragement. In the Encouragement group ranks 1, 3, 5, 7, 9 were shifted to Goal conditions and the even ranks remained as before. The experiment was conducted in the same manner as that described in the first experiment together with the changes as outlined; it was commenced on the Monday following the end of the 10th trial of the first experiment. It should be recalled that there were now six groups, based on the original Control, Encouragement, and Goal groups.

It will sometimes be convenient to make use of the following symbols:

5-C	Five subjects under Control conditions for twenty trials.
4-CG	Four subjects: ten trials Control, ten trials Goal.
5-EG	Five subjects: ten trials Encouragement, ten trials Goal.
4-E	Four subjects: twenty trials Encouragement.
5-G	Five subjects: twenty trials Goal.
4-GE	Four subjects: ten trials Goal, ten trials Encouragement.

(iii) *Results*

As we are now dealing with six new groups, we shall treat the data in terms of the full twenty trials. The results, employing the mean scores for each of the four 5-day periods for each group (the twenty trials were given from Mondays to Fridays for four consecutive weeks) are presented graphically in Fig. 2. (The smoothing procedure employed in plotting the graph did not do violence to the essential pattern of the results.)

As can be seen from Fig. 2 (ignoring for the present the retention score), several fairly clear patterns emerged. Of the two Control groups, 5-C remained more or less unchanged during the 4-weekly periods given, but 4-CG improved considerably when shifted to Goal in the third week, and also exceeded by the fourth week the performance of group 4-E, which remained under Encouragement conditions for twenty trials. The Encouragement group shifted to Goal (5-EG) obtained a higher mean score than 5-C, 4-CG, and 4-E, but remained well below both Goal groups. Both Goal (5-G) and Goal shifted to Encouragement (4-GE) groups remained almost identical throughout except for a slight drop in the scores of those in 4-GE when the shift was made. Ranking the groups according to performance scores for trials 11–20 we had the following order from best to worst:

(1) Goal (5-G).
(2) Goal shifted to Encouragement (4-GE).
(3) Encouragement shifted to Goal (5-EG).
(4) Control shifted to Goal (4-CG).
(5) Encouragement (4-E).
(6) Control (5-C).

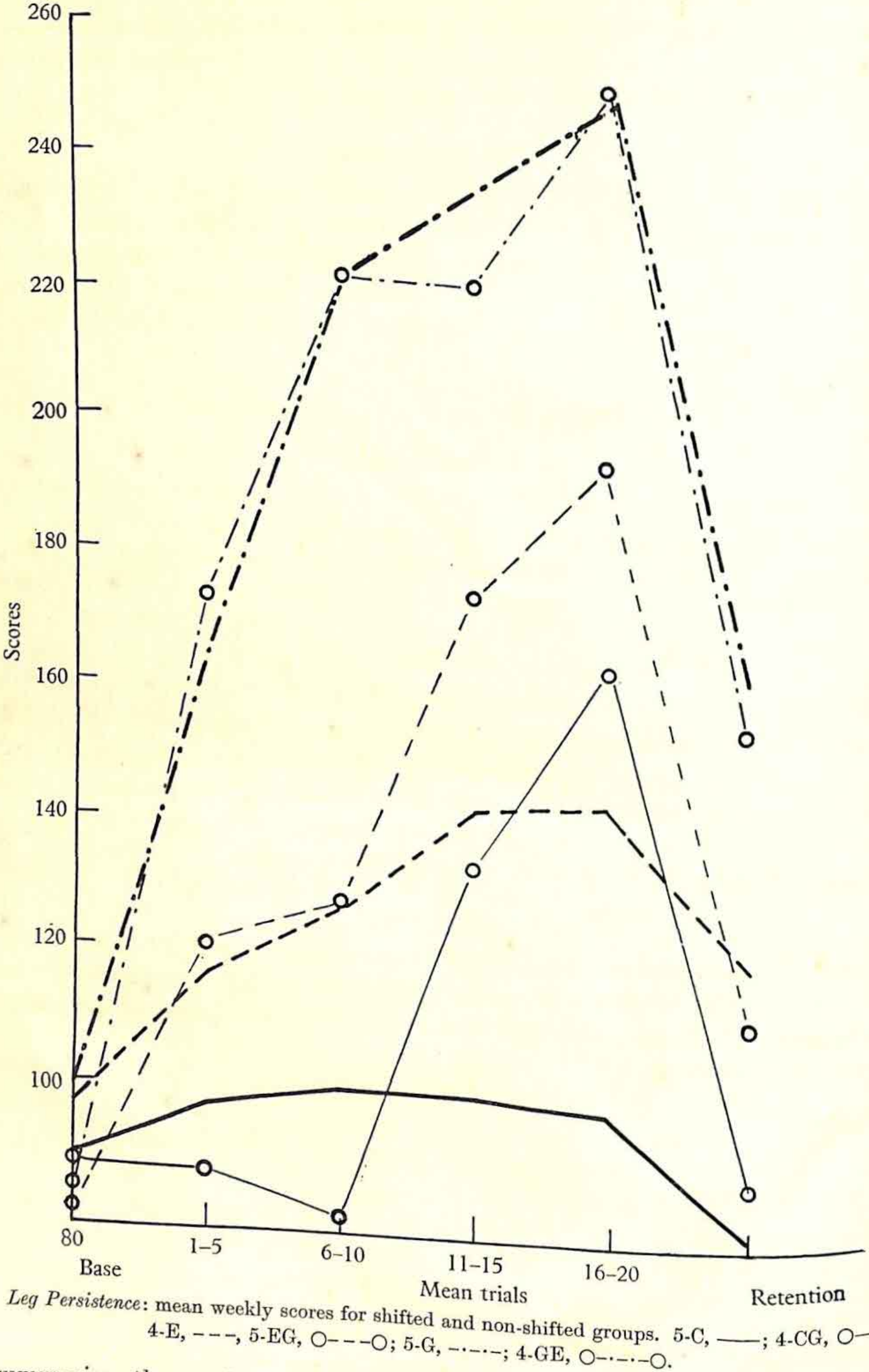


Fig. 2. *Leg Persistence*: mean weekly scores for shifted and non-shifted groups. 5-C, ——; 4-CG, O——O; 4-E, – – –, 5-EG, O– – –O; 5-G, –·–·–; 4-GE, O–·–·–O.

To summarize, the results suggested that both (*a*) type of incentive, and (*b*) prior experience influenced the magnitude of the response. Thus an initially superior incentive showed its effect in superior performance even when replaced by an incentive known to be less effective. (Whether this effect would wear off after a time remained a question for future experimentation.) If, however, a less effective incentive was followed by a superior incentive, the improvement in performance was not as great as that found when subjects began the task under superior conditions of motivation.

(iv) *Conclusions*

The data strengthened the view that the previously obtained results (Experiment One) were not chance findings and that imbecile subjects respond in an orderly manner to incentive conditions. With imbecile subjects, it appeared that *initial* levels of performance were important determinants of subsequent performance. Improvement in performance appeared to be a function not only of the incentives employed but of their sequence.

VI. LEG PERSISTENCE: EXPERIMENT THREE

(i) *Introduction*

To provide further experimental data on the effect of the incentives on subsequent performance the test was again given twice to each subject *under the original matching conditions* after an 8-week interval without practice. The scores obtained are referred to here as 'Retention Scores' by analogy with the retention scores obtained in memory experiments.

(ii) *Method*

After an 8-week interval the subjects were again given the Leg Persistence test under the original Control conditions—that is, a subject was told, when seated, to hold his leg above a chair placed opposite him as long as he could. The subjects were given two trials on subsequent days; the higher of the two scores constituted the measure of retention. During the 8-week interval all but three subjects participated in another series of experiments involving incentives but on an entirely different task.

(iii) *Results*

The retention scores for the six groups are presented graphically in Fig. 2. The results can be summarized in terms of mean scores for each group. They are presented in Table 2 and compared with the mean scores for trials 1–10 of the first experiment and trials 11–20 of the second experiment.

Table 2

Group	Mean score trials 1–10	Mean score trials 11–20	Retention
5-G	193·6	241·3	162·4
4-GE	197·3	234·9	154·3
4-E	121·3	143·3	120·3
5-EG	124·6	184·6	112·0
5-C	99·2	97·7	87·8
4-CG	85·4	148·6	81·0

On examination of the retention scores a curious result was observed. There were large differences between the best group (5-G) and the worst (4-CG). The rank order of the groups did not, however, correspond closely to their scores during trials 11–20. Instead it was the incentives under which they worked during trials 1–10 which seemed to determine the magnitude of their retention scores. If we combine the groups according to the original division into Control, Encouragement and Goal groups the retention scores obtained by each correspond to the order in which the three original groups responded. This is shown in Table 3.

Table 3. *Scores on trials* 1–10 *and on retention for Control, Encouragement and Goal groups*

Group	Mean score trials 1–10	Retention[1]
Goal	195·5	158·4
Encouragement	123·0	116·2
Control	92·3	84·4

Table 4

Group	Retention (1)	Retention (2)
5-G	152·0	147·0
4-GE	144·3	127·5
4-E	120·3	91·5
5-EG	107·4	97·0
5-C	78·0	60·2
4-CG	86·3	64·5

The combined scores are also presented graphically in Fig. 1 (p. 281). The groups which were most alike during the first ten trials remained most alike after an interval of 8 weeks, despite the fact that during trials 11–20 the Control and Encouragement subjects shifted to Goal conditions improved considerably above the level of their counterparts, i.e. those who remained on Control and Encouragement conditions for the full twenty trials.

In order to examine the possibility that these results might be artefacts of the measure used (the higher of the two retention scores), it seemed profitable to examine the two scores separately. The results are summarized in mean scores in Table 4.

The general pattern of the results for either of the retention scores was the same as when the higher of the two scores was used. All mean scores declined on the second retention trial and twenty of the twenty-seven subjects obtained their highest score on the first retention trial.

(iv) *Discussion*

The results again seemed to emphasize the significance of initial achievement levels. Just why the Encouragement and Control groups shifted to Goal conditions (groups 5-EG and 4-CG) did not obtain retention scores which were superior to their non-shifted counterparts we do not know. Examination of the extensive, but by no means exhaustive, reviews of the literature on retention did not reveal an analogous experiment which could have served as a basis for comparison.

VII. CONCLUSIONS

Within the limits imposed by the experimental task employed to reveal some effects of incentives on the behaviour of imbecile subjects, the results led us to accept the view that imbeciles were capable of responding to incentives. Furthermore, as suggested by the first experiment, our subjects responded differentially to incentive conditions such as working towards a goal and encouragement in a manner similar to that noted amongst normal subjects in other investigations.

---

[1] An analysis of variance computed on the basis of the combined groups of the original matching and the retention scores gave an F ratio for the between groups variance of 5·102 significant at less than 0·05 level of confidence for 2 and 16 degrees of freedom.

The persistence test used appeared well suited to reveal these aspects of behaviour. It presented no special problems relating to factors such as prior experience, and scores obtained were largely unrelated to tested intelligence. The results could be objectively scored and were shown to be reliable. As it was a tedious task, devoid of any intrinsic interest, improvement was probably more a function of extrinsic motivation alone than would be the case when subjects were doing something they enjoyed. None of our subjects reported liking the task, but it seemed quite clear that marked improvement in performance was a result of very strenuous effort. The effort, as manifested in performance, was shown to bear a direct relationship to the nature of the incentive offered.

The results of Experiments 2 and 3 strongly support the view that, with imbecile subjects at least, initial levels of achievement are critical determinants of subsequent performance. Both the kind of incentive and the sequence in which it was given appeared to be more significant as factors in improvement than the actual number of trials. However, because of the very small numbers involved in the groups of the later experiments, and the lack of prior experimental work in this area which could serve as a basis for comparison, conclusions must be considered tentative and further discussion must await confirmation of the results.

What emerged most strikingly was that the imbeciles of this sample were 'good' subjects. The results of this exploratory study (involving 648 observations and unaffected by any absenteeism) followed an orderly and consistent pattern and permitted a certain amount of generalization. We gained what appeared to be a very useful basis for hypotheses and for planning further experimental work.

REFERENCES

ALDRICH, C. G. & DOLL, E. A. (1931). Problem solving among idiots. *J. Comp. Psychol.* XII, 137–69.

DOLL, E. A. The feebleminded child. In L. Carmichael (Ed.) (1946), *Manual of Child Psychology*. New York, Wiley.

EDWARDS, A. L. (1950). *Experimental Design in Psychological Research*. New York: Rinehart.

GARDNER, L. P. (1945). The learning of low grade aments. *Amer. J. Ment. Def.* L, 59–80.

KALLMAN, F. J. (1951). Recent progress in relation to the genetic aspects of mental deficiency. *Amer. J. Ment. Def.* LVI, 375–81.

LEWIS, E. O. (1929). Report of an investigation into the incidence of mental defect in six areas, 1925–1927. *Report of the Mental Deficiency Committee, Part IV*. London: H.M.S.O.

LEWIS, E. O. (1948). Mental deficiency and social medicine. *J. Ment. Sci.* XCIV, 258–65.

MACE, C. A. (1935). *Incentives: Some Experimental Studies*. London: H.M.S.O.

TREDGOLD, A. F. (1947). *A Textbook of Mental Deficiency* (7th ed.). Baltimore: Willlams & Wilkins.

(*Manuscript received* 25 *January* 1954)

CHILDREN'S REACTIONS TO BIVARIANT STIMULI

By IAN M. L. HUNTER

Department of Psychology, Edinburgh University

I. *Introductory* (pp. 288–289). II. *Experimental* (pp. 289–291): (i) *Subjects and apparatus* (p. 289–290); (ii) *Procedure* (pp. 290–291); (iii) *Results* (p. 291). III. *Discussion* (pp. 291–293). IV. *Summary* (p. 293). *References* (pp. 293).

I. INTRODUCTORY

One of the major functions of a theory is the parsimonious summarizing of facts. This paper attempts to show that the facts of generalization behaviour as shown by children in the transposition situation may be regarded as special instances of the generalization which occurs in the wider bivariant stimulus situation. This latter situation arises when the stimuli, between which the child must discriminate, offer two differential stimulus components, e.g. there may be both a colour and a form difference as when one stimulus is red and round while the other is green and square. It is the situation made familiar by Descoeudres (1914) and later workers in the investigation of the dominance of colour over number, of form over position, etc. It is basic to conditional and pattern discrimination and is generally representative of those complex situations which are variously described as setting multiple-sign, higher order generalization, or ambivalent-cue problems. A brief review of these situations is given by Harlow (1951).

Suppose a young child is repeatedly presented with two visual stimuli which differ in size and is trained to react always, say, to the smaller. If he is then faced with two new visual stimuli which also differ in size, the child may, under appropriate conditions, transpose, i.e. react to the smaller of the two test stimuli in the same way as he formerly reacted to the smaller member of the training pair. It seems fairly certain (Hunter, 1952) that the frequency of such transposition is an increasing function of: (1) increasing similarity of the two training stimuli to each other in their non-size stimulus characteristics and, within the limits of the discrimination threshold, in their size characteristics; (2) increasing similarity of all the stimulus characteristics of the test pair to those of the training pair; (3) within limits, increasing amounts of practice on the training pair; and (4) increasing mental age of the child.

The first fact above, referring as it does to the conditions favouring abstraction of relational stimulus characteristics, must enter explicitly into any theoretical formulation of transposition phenomena. But, implicit in the writer's previous discussion of the remaining facts (Hunter, 1952, 1953) is the suggestion that they might be explained in terms of two theoretical assumptions. The first assumption (of generalization-by-similarity) is that the amount of generalization shown to any test situation is a direct function of the similarity of that situation to the training situation. Similarity is thought of as occurring along stimulus dimensions which are constructed in terms of either the physical or response properties of the stimulus (or of response properties alone since, on analysis, the former properties reduce without remainder to the latter). The second assumption is that the transposition situation is a special case of the bivariant stimulus situation. It is supposed that, during training, the child learns to react simultaneously to two components of the visual situation which are, logically at least, separable. One component is absolute and involves the specific stimulus characteristics of the entire training situation, including the training stimuli themselves; the other is the relation of 'larger than' or 'smaller than' which, being qualitative, is present or absent in all-or-none fashion.

According to those two assumptions, the frequency of transposition to a test pair which possesses the relational component (i.e. the members of which differ discriminably in size) will depend on the similarity of the absolute characteristics of that pair to those of the training pair. The test pair most similar to the training pair is, of course, the training pair itself which may be thought of as eliciting transposition with the greatest possible frequency. At the other extreme lies some test pair which, absolutely, is so dissimilar to the training pair that the child will fail to react consistently to it in any way whatever. The effects of practice and of mental age on transposition are explained simply by their permitting generalization to take place to test pairs whose absolute characteristics are increasingly dissimilar to those of the training pair. To illustrate, consider generalization at two different stages of training, one shortly after the discriminative reaction has been established and the other after overtraining. At the first stage, reaction to the test pair is haphazard, i.e. a change in the absolute component brings about discrimination breakdown. At the second stage, the same change in the absolute component does not produce breakdown and transposition appears.

The above way of restating the facts, while admissible, is hardly parsimonious unless some general statement can be made about the interaction between the two hypothesized stimulus components and the influence which practice and mental age have on that interaction. The experimental part of this paper is directed toward providing such a statement.

The validity of treating transposition in terms of reactions to a bivariant stimulus situation cannot, however, be checked directly within the transposition situation itself since the two stimulus components are not, there, independently manipulable. Specifically, while the absolute component can be altered without disturbing the relative, the latter cannot be removed without tampering with the former. In order to study each of the two components, a bivariant stimulus situation must be devised which meets two requirements. First, each component must be independently variable with regard to the other. Secondly, one component must not be so very much more readily reacted to than the other that only the dominant one enters effectively into the picture. The discrimination situation to be described below meets both requirements.

II. EXPERIMENTAL

(i) *Subjects and apparatus*

The subjects were twenty-seven nursery school children (twelve boys, fifteen girls) with an age range of 38–60 months and an average age of 50·2 months.

The training and test pairs each comprised two 6 × 6 inch stimulus cards presented, side by side, on an apparatus so that, immediately beneath each, was a small drawer. This simple presentation apparatus has already been described (Hunter, 1952). The child discriminated between the two cards by opening the drawer beneath one of them. During training, there was always a raisin in the drawer beneath the positive card and, during testing, there was a raisin in each drawer. The cards, cut from stiff white cardboard, were painted mat black to leave a white geometrical figure exposed in the centre. These figures were bivariant being made up of varying sizes of upright equilateral triangles broken by alternating black-and-white horizontal stripes of varying width. There were three sizes of triangle—'large', 'medium', and 'small'—with sides 11·3, 9·5, and 8·0 cm. long, respectively. The area of the large triangle was, therefore, 1·4 times that of the medium which, in turn, was 1·4 times that of the small. Likewise, there were three widths of stripe—'thick', 'medium', and 'thin'—which measured 0·4, 0·3 and 0·2 cm., respectively. The black and white stripes were of the same width on every card so that any difference in the total brightness of two cards was due entirely to the different sizes of the triangles on them. Since linear size and total brightness co-varied and since their separate contributions to discrimination were not being analysed, these two characteristics were simply treated as constituting the single characteristic of 'size'.

Eight stimulus cards were prepared, each bearing one of the nine possible size-width combinations (the medium-medium combination was not required). These eight combinations were: large-thick, large-medium, large-thin, medium-thick, medium-thin, small-thick, small-medium, and small-thin. Because the altitudes of the triangles were not all exact multiples of the stripe widths, incomplete stripes at either apex or base might have furnished unwanted differential features. To avoid this, the altitudes were altered very slightly to make them exact multiples of the particular stripe width with apex and base stripes both white. These eight cards were split into two training pairs and two test pairs. One training pair was: small-thin (+) *v.* large-thick (−). The other training pair was: small-thick (+) *v.* large-thin (−). Both pairs provided the two components of differential size and differential width. The two test pairs, on the other hand, provided only one component each, the differential features of the other being removed. The test pair which provided only the differential size component was: large-medium *v.* small-medium. The other test pair, providing only the differential width component, was: medium-thick *v.* medium-thin.

(ii) *Procedure*

On being brought into the experimental room, the child was seated in front of the presentation apparatus, shown the cards and the drawers beneath them, and allowed to take a raisin from one of the drawers. Training was then begun. The training pair was presented and the child told, by word and gesture, to 'look at both of the cards', then to 'point to the one that has the raisin', and then to 'open it'. If the drawer beneath the positive card was opened, the child found a raisin which he could eat. If the 'wrong' drawer was opened, the child then found that he could not open the other and was told that he must open the 'right one' first otherwise he would not get the raisin. This was the first trial. For the second trial, the positions of the cards were changed and the above procedure repeated. Thereafter, the two training cards, their relative positions being changed in random order, were presented repeatedly until the child reached the arbitrary learning criterion of five consecutive trials correct. With the older children, it was not necessary to repeat the three separate commands on all trials; it sufficed to say 'Choose'. But unless all three directions were consistently given to the younger children, they began to react to the drawers without looking at the cards at all. Preliminary observations showed that it was the task of manipulating some part of the apparatus other than the stimulus cards themselves which made this discrimination arrangement too difficult for very young children. Fourteen children were trained with one training pair, and thirteen with the other. All children were tested with both test pairs.

When the child had reached the learning criterion, he was given three consecutive presentations of one or other of the test pairs. Training was then resumed until two consecutive training trials were correct, whereupon three trials were given with the remaining test pair. Thereafter, training trials were continued with groups of three test trials interspersed irregularly among them. The child was regarded as showing generalization to a particular test pair when he chose the same member of that pair on three consecutive trials regardless of position. The experiment was concluded when the child generalized to both test pairs. Only one subject refused to continue before reaching the end of the experiment.

(iii) *Results*

On the average, the children gained the learning criterion in seventeen trials and, after a further twenty-two training trials (this figure does not include test trials), showed bidimensional generalization, i.e. generalization to both the size and width components when each is presented in isolation from the other. With six subjects, a further twenty training trials in no way altered this bidimensional generalization, suggesting that it was a stable end-stage. (These averages do not take account of the one child who finished before the experiment was completed. He gained the learning criterion in seventeen trials and left the situation after a further six training trials at which time he was generalizing in terms of size but not of width. However, there was no reason to doubt that, like the other children, he would have generalized to both components had he continued. In presenting the results below, it will be assumed that he did continue and did show bidimensional generalization.) Once the child had reached the learning criterion, few if any mistakes were made on subsequent training trials. When, during testing, one card was not consistently chosen, the usual reaction was to position, e.g. always to the left side. Choices against the direction of training, in the sense of having been trained to choose the small triangle and then choosing the large triangle in a test trial, often appeared but never consistently.

In the test trials, four different types of generalization were possible, i.e. generalization to neither component (discrimination breakdown), generalization to both components (bidimensional generalization), generalization to the size component only, and generalization to the width component only. Thus, individual performances could differ from each other in terms of the number of types of generalization shown and in the order in which they made their appearance as training proceeded. In terms of such differences, the children fell into four groups. (1) Nine children (age range, 49–60 months; mean age, 57 months) showed only bidimensional generalization. (2) Seven children (age range, 39–53 months; mean age, 47 months) showed, first, generalization to size, then bidimensional generalization. (3) Four children (age range, 46–57 months; mean age, 50 months) showed, first, generalization to neither component, then bidimensional generalization. (4) Seven children (age range, 38–58 months; mean age, 47 months) showed, first, generalization to neither component, then generalization to size, and lastly bidimensional generalization.

III. DISCUSSION

The first general trend indicated by the results is that each child passes through the same generalization sequence in which, as training proceeds, there can, in turn, be removed without discrimination breakdown, first, *neither* of the two stimulus components, secondly, the width component, and lastly, *either* the size *or* the width component. It may be that the above four groups result from the children's passing through genuinely different generalization stages, but it seems more likely that the performances of those children not showing all three stages can be accounted for in terms of the stages succeeding each other so rapidly that either, or both, of the first two escape detection. The experimental method employed could hardly be expected to sample every generalization stage through which every individual passes so that the failure to demonstrate a particular stage in any child does not at all mean that he does not pass through it. The absence of the first two stages in some subjects affirms that bidimensional generalization is not merely due to increasing familiarity with the test stimuli. On a more general plane still, these sequences of generalization types illustrate the basic developmental process of 'perceptual articulation', the process which Werner describes as 'the development from a diffuse perceptual organization characterized primarily by "qualities-of-the-whole" into an organization in which the essential feature is a decisiveness of parts standing in clear relationship' (Werner, 1940 pp. 111–12). Perceptual articulation plays a part in many discrimination situations, and Murphy and Hochberg go so far as to claim that 'progression from (*a*) homogeneous through (*b*) differentiated to (*c*) integrated characterizes, in principle, not only the stages through which the perceptual development of the maturing child passes, but also (with important qualifications...) the course of every perceptual process, regardless of age.' (Murphy & Hochberg, 1951, p. 334).

The second general trend indicated by the above results is that perceptual articulation is facilitated not only by increasing amounts of training but also by increasing chronological age. The children showing bidimensional generalization straight away are an average of nine months older than those who show, first, any of the more primitive types of generalization. This age difference is, however, no more than suggestive since there are too few subjects to permit statistical analysis. But the age difference is in the right direction (see below) and the writer believes that it would be worth seeing if it became larger if mental age were considered rather than chronological age.

The results also show that, with the particular stimuli used, the size characteristic is more readily reacted to than the width characteristic. There is not a single case of a child's generalizing to the width component before being able to generalize to the size component.

It is now time to return to the child's generalization behaviour in the transposition situation. As suggested in the Introduction, the training stimuli in that situation may be regarded as bivariant, as possessing both absolute and relational components. By analogy with the above suggested trends, with younger children and small amounts of training, the stimuli will give rise to a diffuse perceptual organization, and the alteration of either component will lead to discrimination breakdown, i.e. transposition will not occur except with test pairs which preserve the relational component and are, in their absolute characteristics, very similar to the training pair. On the other hand, with older children and large amounts of training, the perceptual organization resulting from the two components will become progresively more articulated and transposition will appear to test pairs increasingly dissimilar to the training pair.

Another effect not infrequently found in the transposition situation is that, consequent on the increased frequency of transposition with increasing amounts of training, further training leads to a decrease in transposition and a choice of the positively trained stimulus or of stimuli similar to it. This effect, too, can be viewed in terms of perceptual articulation if a third stimulus component is admitted which comprises the absolute characteristics of the positive stimulus. Generalization in terms of this component is, of course, incompatible with generalization in terms of the relational component and less frequent transposition will result.

IV. SUMMARY

The generalization behaviour of twenty-seven children, of average age 50·2 months, is investigated in a bivariant stimulus situation. It is found that, when the discriminative reaction is newly established, removal of either stimulus component produces discrimination breakdown but that, with further training, generalization occurs to test situations which provide either component in the absence of the other. Results are fully consistent with the hypothesis that increasing amounts of 'perceptual articulation' occur with increasing amounts of training and, further, occur more readily the older the child. The concept of 'perceptual articulation' is used to interpret the generalization behaviour previously observed in the transposition situation. It is argued that the transposition situation may profitably be regarded as a special case of the more inclusive bivariant stimulus situation.

The writer wishes to express his sincere thanks for their cooperation to: the Matron of Broomhead Day Nursery, Dunfermline; and the Headmistress and Staff of Moray House Nursery School, Edinburgh.

REFERENCES

DESCOEUDRES, A. (1914). Couleur, forme ou nombre? *Arch. de Psychol.* xiv, 305–41.

HARLOW, H. F. (1951). Thinking. In H. Helson, *Theoretical Foundations of Psychology*, pp. 452–505. New York: Van Nostrand.

HUNTER, I. M. L. (1952). An experimental investigation of the absolute and relative theories of transposition behaviour in children. *Brit. J. Psychol.* XLIII, 113–28.

HUNTER, I. M. L. (1953). Reply to Professor Grice. *Brit. J. Psychol.* XLIV, 161–2.

MURPHY, G. & HOCHBERG, J. (1951). Perceptual development: some tentative hypotheses. *Psychol. Rev.* LVIII, 332–49.

WERNER, H. (1940). *Comparative Psychology of Mental Development.* New York: Harper Bros.

(*Manuscript received 8 February* 1954)

THE EFFECT OF HUNGER UPON DRINKING PATTERNS IN THE RAT

By ALLEN D. CALVIN and RICHARD A. BEHAN
Michigan State College

I. *Introduction* (pp. 294–295). II. *Method* (p. 295). III. *Results* (pp. 295–297). IV. *Discussion* (pp. 297–298). V. *Summary and Conclusions* (p. 298). *References* (p. 298).

I. Introduction

Although animals have been used extensively in behavioural studies involving thirst and hunger, attempts to study the actual interaction of these two complex drive mechanisms seldom appear in the literature. Siegel and his associates have published three studies in this area. Siegel & Stuckey (1946) investigated the relationship between water intake and food intake in the rat. Although their study was primarily concerned with the validity of Richter's hypothesis that there is a direct relationship between skin area and water intake in the rat, they reported (p. 273 n. 3) that their rats drank 29% less water when they fasted than when they were on *ad libitum* feeding. In another study the same authors (1947) found that the food and water intake of rats showed a marked correspondence when the animals were given food and water *ad libitum*. Siegel and Talantis (1950) found that depriving rats of food did not alter the relationship between water intake and hours of water deprivation.

Finger & Reid (1952) and Strominger (1947), working with rats, report that when rats are deprived of food, they voluntarily restrict their water intake.

Some earlier studies also attempted to investigate the interaction of thirst and hunger. Gregerson (1932), working with dogs, found that fasting usually reduced the 24 h. intake of water to one-fourth normal or less and that almost all of the water consumed by his experimental animals was drunk within 2–5 h. after feeding regardless of the time when food was given. Gregerson believes this is attributable to a temporary state of dehydration due to withdrawal of water from the body for the secretion of digestive juices. Kleitman (1927) reported that starving dogs showed a decrease of 65–80% in total water intake. In another study, using dogs, Robinson & Adolph (1943, p. 41) reported; 'It is shown above that dogs which were fed many hours ago drink little water. This fact has long been recognized in that dogs drink only one-tenth to one-third as much water upon days when food is withheld as upon days of usual feeding.' They further report on p. 43: 'It was found that water was usually drunk a few minutes after dry food was eaten. Since the reverse rarely occurred, some support could be deduced for the notion that eating sets off drinking.... The important point here is, however, that if the dog's five meals a day are stopped by deprivation of food, drinking that ordinarily follows each meal goes on just the same. The amounts drunk are reduced, but the same number of draughts are taken per day. Hence the pattern of drinking is demonstrated to exist independently of eating, though ordinarily the two work together in one sequence.'

The present experiment was undertaken to study quantitatively the effect of hunger and thirst upon drinking behaviour in the rat under varied conditions of food and water deprivation.

II. METHOD

Forty naïve albino rats, twenty males and twenty females, 90–120 days of age at the beginning of the experiment, were used in the present study. The experiment was run in two successive parts, each of which utilized ten males and ten females. All animals were on an *ad libitum* food and water schedule up to the beginning of the experiment. The drinking tubes employed were standard chemical laboratory cylinders, graduated to 1 ml. Accurate readings could be made to ½ ml.

During part I of the experiment ten animals, five males and five females, were placed in individual feeding cages at 8 a.m. and deprived of food and water for 24 h. At the end of the deprivation period water was introduced. The drinking tubes were read after the following time periods: 2, 5 and 10 min., and 1, 12 and 24 h. Five of the ten animals were also given 35 g. of Purina Dog Checkers at the time that water was introduced.

The other ten animals, five males and five females, used in part I were treated exactly as the first ten except that they were started on the experiment at 11 p.m.

Part II of the experiment was a duplication of part I except that the animals were placed on a limited feeding schedule for 8 days prior to the 24 h. deprivation period. During this 8 day period the animals were each given 9 g. of Purina Dog Checkers per day, and water was available at all times.

The readings on the night animals of group II were taken 17 days after the last of the readings were made on the animals of group I.

III. RESULTS[1]

Separate analyses of co-variance using body-weight as the correlated variable were run for each time of reading, since the variances for the different times of reading were not homogeneous. The analyses were run with respect to the following variables: sex, feeding schedule (FS) *v.* no feeding schedule (no FS), food and water *v.* water alone and night *v.* day.

At the 2 min. readings there were no differences with respect to any of these variables. At the 5 min. readings the water intake of the food and water animals (the FW group) was significantly greater (5% level of confidence) than the water intake of the water only animals (the W group). Differences with respect to the other variables were not significant.

At the 10 min. and 1, 12 and 24 h. readings the water intake of the FW group was significantly greater (1% level of confidence) than the water intake of the W group. The F-values were 8·36, 46·19, 146·45 and 265·40 respectively. None of the other variables achieved a significant F-value, and the F-values of these other variables were not ordered as were those for the FW and W groups.

---

[1] After the present study had been completed and the manuscript prepared for publication, a study by Verplanck & Hayes (1953) appeared. Since the present study presents some findings which correspond with the findings for the O*w* (21½ hr. of food and water deprivation followed by water) and O*fw* (21½ hr. of food and water deprivation followed by food and water) groups of the study of Verplanck & Hayes we prepared a chart to compare our 1 hr. readings after 24 hr. of food and water deprivation with theirs. The difference in deprivation time of the two studies prevents strict comparability of results.

	V. & H. mean	C. & B. mean
O*w*	4·18	3·9
O*fw*	10·05	11·0
w/fw	0·42	0·36

Thus, the values we obtained correspond quite closely to those obtained by Verplanck & Hayes.

Table 1 compares the mean water intake for the W group and the FW group for the different times of reading. It is seen in Fig. 1 that water intake is a negatively accelerated function of the time since water was made available, which corresponds with the findings of Stellar & Hill (1952). The asymptote for the FW group is considerably higher than that for the W group.

Fifty per cent of all the animals drank immediately regardless of whether food was introduced with the water.

Pearson product-moment correlation coefficients were computed separately for the W and FW groups. Of the six pairs of correlation coefficients so obtained only the pair for the 2 min. reading was significant. Since the difference between any pair was less than the standard error of either member of the pair, the correlations were re-computed for the different times of reading with the data for the W and FW groups combined, thus increasing the D.F. to 38.

Table 1. *Mean water intake at the six times of reading for the water group and the food and water group*

Time of reading	Water group	Food and water group	Probability of significant difference
2 min.	1·6	1·2	—
5	2·1	2·8	<0·05
10	2·2	4·3	<0·01
1 h.	3·9	11·0	<0·01
12	7·1	32·3	<0·01
24	8·1	37·6	<0·01

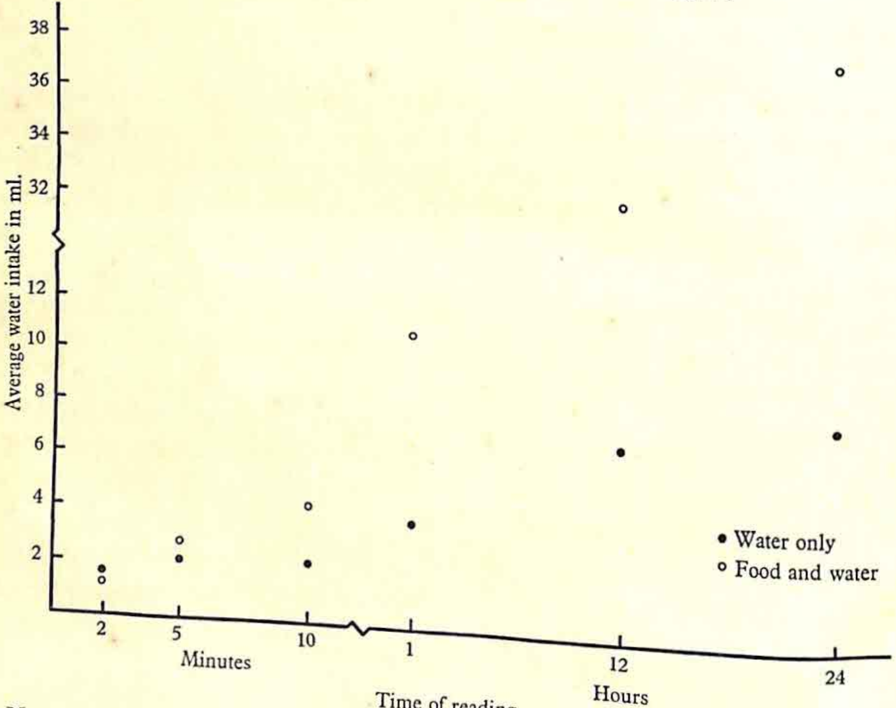


Fig. 1. Mean water intake at the six times of reading for the water group and the food and water group.

The correlation coefficients between water intake and body weight were as follows. For the 2, 5 and 10 min., and 1, 12 and 24 h. readings these were −0·44, −0·34, −0·31, 0·17, 0·03 and 0·04 respectively. The first is significant beyond the 1% level of confidence, the second and third beyond the 5% level of confidence, and the last three are not significant.

The corresponding values of η for the regression of water intake on body weight are 0·60, 0·55, 0·52, 0·38, 0·29 and 0·19. The first four are significant beyond the 1% level of confidence, and the last two are not significant.

Deviation from linearity is significant in the case of the 2, 5 and 10 min. readings (1% level of confidence), and in the case of the reading for 1 h. (5% level of confidence). Deviation from linearity is not significant in the case of the 12 and 24 h. readings.

IV. DISCUSSION

The inverse curvilinear relationship between water intake and body weight found in the present study was a source of considerable surprise since Siegel & Stuckey (1946) found a low positive, but not significant, correlation between body weight and water intake when the food deprivation period was increasing from 0 to 23 h. and 45 min. It is to be remembered, however, that the animals of the present study had already been deprived of food and water for 24 h. before the drinking tubes were introduced.

Of further interest is the finding that the magnitude of the correlation between body-weight and water intake decreases as the amount of time that water is available increases, regardless of the presence or absence of food, up to 24 h. which was the limit in the present study.

The FW group drank a mean of 37·6 ml. of water over a 24 h. period as compared with 8·1 ml. for the W group. Siegel & Stuckey (1947) found that the average water intake for a 24 h. period with animals on an *ad libitum* food diet was 32·6 ml. Both of the values (8·1 and 37·6) of the present study are beyond the range of the true mean of the Siegel & Stuckey study. This may of course be attributed to the difference in experimental procedures mentioned earlier.

The fact that the present study gives no evidence of a differential water intake between night and day when animals are deprived stands in direct contradiction to results presented by Siegel & Stuckey (1947), Stellar & Hill (1952), Young, *et al.* (1952). We can offer no explanation of this contradiction, but can only point out that our rats were deprived of food and water for 24 h. before the test situation which may well have altered the drinking behaviour which was observed by the other investigators. Young, *et al.* (1952, p. 85) made a similar observation: 'From our records it is clear that a 6 hour deprivation from 6 a.m. till noon would have little effect upon the water intake of our animals inasmuch as they normally drink little during these hours. But an equal period of deprivation during the night hours might have a very different effect upon drinking behaviour, especially if the thirst drive were tested at night.'

Of equal interest was our finding that there were no significant differences between the FS and the noFS groups. Restricting the food intake of the animals of part II for 8 days prior to depriving them of both food and water for 24 h. had no effect on their water intake, for when their behaviour was compared with that of the part I animals, which had been on *ad libitum* feeding prior to their 24 h. period of water and food deprivation, no significant differences between the groups were discovered.

Our W group's drinking only about one-quarter as much water as the FW group is in agreement with the results of other investigators (Gregerson, 1932; Kleitman, 1927; Siegel & Stuckey, 1946). The finding of the present study that there is reduced water intake in the W group tends to support the hypothesis that eating sets off drinking. Our finding that there is reduced water intake coupled with the findings of Stellar & Hill (1952) that rats drink at a constant rate or do not drink at all, allows the deduction that rats do not take the same number of laps when food is available and when it is not. Robinson & Adolph (1943) reported that while dogs took in less water when food was not available they still took the same number of draughts. Since Stellar & Hill counted laps and Robinson & Adolph counted draughts we are not really in a position to make a comparison and the difference may be explained simply on the basis of the variable considered.

V. Summary and conclusions

Water intake of forty albino rats was measured after 2, 5 and 10 min., and 1, 12 and 24 h. under two conditions of food-deprivation. Group I consisted of twenty animals which were deprived of food and water for 24 h. Group II consisted of twenty animals which were put on a 9 g. per day food schedule for 8 days and then deprived of all food and water for 24 h. Half of the animals of each group were started during the day and half at night; half of each of these groups had food available during the period of water intake measurement and half did not.

Quantitative measures of the effect of hunger and thirst upon drinking behaviour in the rat under varied conditions of food and water deprivation are presented.

References

Finger, F. W. & Reid, L. S. (1952). The effect of water deprivation and subsequent satiation upon general activity in the rat. *J. Comp. Physiol. Psychol.* XLV, 368–72.

Gregerson, M. I. (1932). Studies on the regulation of water intake. II. Conditions affecting the daily water intake of dogs as registered continuously by a potometer. *Amer. J. Physiol.* CII, 344–9.

Kleitman, N. (1927). The effect of starvation on the daily consumption of water by the dog. *Amer. J. Physiol.* LXXXI. 336–40.

Robinson, E. A. & Adolph, E. F. (1943). Pattern of normal water drinking in dogs. *Amer. J. Physiol.* CXXXIX, 39–44.

Siegel, P. S. & Stuckey, H. L. (1946). An examination of some factors relating to the voluntary water intake of the rat. *J. Comp. Physiol. Psychol.* XL, 271–4.

Siegel, P. S. & Stuckey, H. L. (1947). The diurnal course of water and food intake in the normal mature rat. *J. Comp. Physiol. Psychol.* XL, 365–70.

Siegel, P. S. & Talantis, B. S. (1950). Water intake as a function of privation interval when food is withheld. *J. Comp. Physiol. Psychol.* XLIII, 62–5.

Stellar, E. & Hill, J. H. (1952). The rat's rate of drinking as a function of water deprivation. *J. Comp. Physiol. Psychol.* XLV, 96–102.

Strominger, J. L. (1947). The relation between water intake and food intake in normal rats and rats with hypothalamic hyperphagia. *Yale J. biol. Med.* XIX, 279–88.

Verplanck, W. S. & Hayes, J. R. (1953). Eating and drinking as a function of maintenance schedule. *J. Comp. Physiol. Psychol.* XLVI, 327–33.

Young, P. T., Heyer, A. W. & Richey, H. W. (1952). Diurnal drinking patterns in the rat. *J. Comp. Physiol. Psychol.* XLV, 80–9.

(*Manuscript received* 16 *February* 1954)

PUBLICATIONS RECENTLY RECEIVED

The Nature of Prejudice. By G. W. ALLPORT. Massachusetts: Addison-Wesley Publishing Co. 1954. Pp vxiii + 537. $5.50.

Prejudice is possibly the most fully studied of all topics in social psychology, and there is no doubt of its great practical importance in this time of awakening nationalism amongst the coloured populations of the world and the anxieties accompanying that awakening in the advanced western peoples. Allport in this extensive review succeeds in demonstrating both these points, and, in these days of fashionable 'fundamental psychology' confined to the laboratory, shows that at least one branch of psychology has something reasonable and useful to say about human affairs and their conduct.

His approach is a model one. He first defines prejudice; a definition which is careful to include all forms of prejudging and which does not restrict itself to the more obvious forms of colour and ideological prejudice. Most of us are prejudiced in one way or another if we accept his definition and, as he implies, even the most liberal person may be violently prejudiced in some ways. The negrophil is often as prejudiced as the negrophobe. Not only are we all prejudiced but without a certain amount of prejudging of persons and issues we could not behave effectively; prejudice is a useful and a normal form of behaviour which can only be regarded as pathological when it leads to maladaptive behaviour and gross mistakes in estimating the nature of the environment.

This view of prejudice at once makes it a matter of character structure and Allport produces ample evidence that this is the case. He deals fully with the ontogenesis of prejudice and reviews fully all the studies on prejudice formation in childhood. The recent material of the Frenkel-Brunswick school occupies an important place in the book, together with the older studies of the personality structure of the prejudiced person. And Allport does not restrict himself; he gives full weight to the clinical and psychoanalytic studies as well as to the more quantitative and rigorous investigations. This catholic approach to the material is one of the most stimulating features of the book, and it is to be hoped that the custom of regarding all evidence as useful in a beginning science will become usual.

The emphasis on the character structure of the prejudiced person implies that the removal of prejudice is less a matter of instruction than of therapy. The later chapters of the book discuss the experiments already performed on the cure of prejudice and, so far as it goes, the evidence supports the need for special quasi-therapeutic techniques. But it is in this field that least has been done, and it is clear that there is a great need for more experiments on methods of destroying prejudice. It might be the case that if we had useful techniques our political tasks in Africa and the East would be made easier.

Allport does not make the common professional mistake of neglecting social and economic causes of prejudice. However, he does show that they operate by affecting the acts of individuals and regards the psychologist's task as done when he has explained the precise manner of their action on the character of the individual. He briefly but adequately reviews the work on the social and economic causes and shows their relation to the character of the prejudiced individual.

There is no better survey of the subject of prejudice, and this book is likely to become a standard text for social psychologists beginning the study of prejudice. It is spoiled a little by equal weight being given to hearsay evidence and to more certain material. However, it is easy to distinguish which is which in the text. Lastly, the book is very readable and can be easily understood by laymen. For this reason it should have a wider audience than the professional psychologist whose technical voice is rarely understood in political and administrative circles.

R. C. ALBINO

Tâche, Réussite et Échec.—Théorie de la Conduite Humaine. By J. NUTTIN. Antwerp: N. V. Standaard-Boekhandel. 1954. Pp. ix + 530. Sewn, 280 Belgian francs; bound, 330 Belgian francs.

The subject of this massive work by Professor Nuttin is the interplay between motivation, past experience and present apprehension in the determination of adaptive behaviour and the development of personality. Its scale matches appropriately the breadth of the theme, and it incorporates a great quantity of experimental work, for the most part hitherto unpublished. The empirical material is sustained by most careful discussion, explicitly based upon the fullest study of contemporary and recent work in relevant fields.

It is Nuttin's contention that no simple hypothesis of temporary need-reduction, even when supported by a conception of secondary motivation, can fully compass the motivational aspects of learning and performance as they are observed in human subjects. He has conducted a series of very ingenious experiments, for the most part concerned with the effects of apparent success and failure upon subsequent recall both of experimental material and of success or failure themselves. These lead him to the conclusion that performance, while it arises within a frame of existing needs, leaves behind it not merely the *reduction* of a need, but also the *persistence* of a need. This later becomes incorporated in a stably evolving system of orientation, at once cognitive and dynamic, in the individual. The need-reduction aspect of successful performance has, thus, far less importance from a broad psychological point of view than that attributed to it by many behaviouristic theorists. What is of much greater interest is the manner in which success or failure contributes to the persistence and evolution of the individual's motivational system. For it is through this effect that fresh ways of regarding the environment, and of developing expectations, significances and potential satisfactions within it, are engendered. And so in turn the means of bringing past experience and established habits to bear in novel adaptive reactions are developed.

No brief review could do justice to the elaborateness with which Professor Nuttin has developed this theme. He has, in effect, made an outstanding effort to weave into a unified theoretical formulation all the manifold considerations and contradictions afforded by such diverse approaches as those of Thorndike, Clark Hull, Tolman, the Gestalt School and Bartlett. The meal is large, rich, varied and most skilfully prepared, if a trifle difficult to digest. One is left a little uncertain, however, whether the basic materials really lend themselves to this kind of *liaison*.

R. C. OLDFIELD

The Structure of Human Personality. By H. J. EYSENCK. London: Methuen. 1953. Pp. 348. 37*s*. 6*d*.

In this book Dr Eysenck gives a thorough and sometimes sharply critical account of a large number of studies, mainly factorial, of various aspects of personality, using rating methods, questionnaires, objective behaviour tests, measurements of physique and physiological measurements, and assessments of interest and attitude. He shows that in the main the results of these studies are compatible with the 'dimensions' or principles of classification which he has set forth in previous work. He also shows that there are fewer differences between the results of studies using different factorial methods than might be thought, especially when second-order analysis is done on oblique primary factors. One would not deny that factorial methods can be of considerable value in enabling us to clarify and improve our classificatory categories. Many, however, would not agree with Dr Eysenck that the value of typologies or systems of classification can be established only by factorial methods.

Dr Eysenck lays stress on the use of the 'hypothetico-deductive method'. But his use of the term 'hypothesis' appears to the reviewer somewhat doubtful. For example, he claims to test 'Jung's hypothesis that the conduct of extraverts is more determined by objects and relations in the external world, while the actions of introverts are more determined by their inner subjective states' (p. 29). He uses the 'level of aspiration' technique, and finds that introverts (represented by dysthymics) aspire higher than they perform, and perform better than they estimate, while the reverse is true of extraverts (represented by hysterics). This is taken as supporting the hypothesis. Now, assuming that one accepts Dr Eysenck's selection of his two groups, the question is whether the hypothesis is a genuine hypothesis in the sense that it could be disproved conclusively by experimental procedure. One might legitimately argue that any experimental procedure such as Dr Eysenck's is not unambiguously relevant to the hypothesis, so that a critical experiment would not be possible. To put it in another way, all that one could conclude from a negative result would be that the behaviour referred to by the terms 'introvert' and 'extravert' does not include the behaviour in the experimental situation. This is the same kind of point as that made by Albino in his recent article in this Journal. It does not, of course, imply that it is not interesting and valuable to endeavour to define or extend the range of classificatory categories.

A few other points might be made. For example, one would like to know how it comes about that rating of obsessional tendencies has a zero-loading on a factor of 'neuroticism' in one study (p. 55), while in another study involving subjects classified as cases of psychopathy, obsession, hysteria, anxiety state or post-traumatic personality change, 'the first component (severity of neurosis) ranks the groups in order from normal to obsessional and psychopathic' (p. 57).

Such questions, however, should not be allowed to detract from the principal merits of the book, which should be read by all interested in the assessment of personality.

D. GRAHAM

The Origins of Psycho-Analysis. Letters to Wilhelm Fliess, Drafts and Notes; 1887–1902. By Sigmund Freud, edited by Marie Bonaparte, Anna Freud and Ernst Kris. London: Imago Publishing Co. Ltd. 1954. Pp. xi+486. 30*s.*

This book contains a selection of intimate letters written by Freud during the period of his greatest originality of thought to Wilhelm Fliess, a biologically minded physician in Berlin, under the influence of a friendship which evidently provided Freud at a critical phase with the inspiration to pursue his momentous researches undeterred by difficulty, discouragement and disappointment. Fliess was an aberrant genius who came to harness the concept of bisexuality to a pseudo-scientific system of periodicity based on female menstruation. His growing extravagance of thought and intolerance of disagreement ultimately led, in 1902, to an alienation of Freud's friendship; but the emotional basis of the friendship had already become undermined by the arduous self-analysis commenced by Freud in 1897. The influence of this self-analysis upon the development of Freud's psycho-analytical theories, e.g. his theory of dreams, is also reflected in the letters—as is the conflict (never wholly resolved) between his ambition to provide a purely psychological explanation of psychopathological phenomena and a lingering desire, encouraged by Fliess' influence, to achieve an explanation in physiological terms. A remarkable wealth of theoretical material is combined in the letters with interesting personal detail; and the drafts accompanying certain letters contain material never embodied in Freud's published works. The book will provide the student of psycho-analysis with a fascinating study; and for the intimate light which it throws upon the manner in which a scientific genius works it is almost unique in literature. W. RONALD D. FAIRBAIRN

Psychoanalysis as Science. By Ernest R. Hilgard, Lawrence S. Kubie and E. Pumpian-Mindlin. Stanford University Press; London: Geoffrey Cumberlege. 1952. Pp. x+17. 34*s.*

Psychoanalysis and the Unity of Science. By Elsie Frenkel-Brunswik. Proceedings of the American Academy of Arts and Sciences, Vol. LXXX, No. 4. 1954. Pp. 271–350. $1.55.

Psychoanalysis as Science consists of five lectures delivered at the Carnegie Institute of Technology, 'to acquaint an audience of scientists from other fields with the basic principles underlying, and the basic evidence for, psychoanalytic concepts'.

In the first two lectures, Dr Hilgard employs his outstanding powers of exposition in summarizing the attempts that have so far been made to subject Freudian hypothesis to experimental tests. The lectures confirm Freud's own prediction that 'the future will probably attribute far greater importance to psycho-analysis as the science of the unconscious than as a therapeutic procedure'. They contain some important material on experimental neurosis in human beings and animals, but no reference is made to attempts to measure the success, or otherwise, of psychoanalysis in the treatment of real life neurotics. On the other hand there is an impressive array of experimental evidence for the validity of the basic Freudian concepts, particularly the defence-mechanisms and dream symbolism. Some of the material was summarized in Sears' 1944 survey, but much of it is more recent. Of special interest is the full account of Keet's 'micro-analysis', originally described in *Psychological Monographs*, 1948.

Dr Hilgard himself is, of course, by no means a Freudian. He sums up his theoretical position by saying, 'I do not care whether we end up believing that Freud was a scientist or a romanticist. If eventually we are able to make science out of materials that he called strongly to our attention, then we shall owe him that historical debt, whatever the verdict may be about his own formulations.'

In the third and fourth lectures Dr Kubie, a Professor of Psychiatry and a practising analyst, questions the value of providing 'pallid laboratory facsimiles' of facts that have already been abundantly proved in real life; and he gives a string of examples—some from everyday experience and others from clinical practice—where, in his view, the Freudian implications are too obvious to reject. One or two of these examples are open to scepticism, but most of them are convincing and striking: they must have considerably startled an audience new to the subject.

In Dr Kubie's view, the function of experiment in psychoanalysis is to advance knowledge, not to convince those who are 'reluctant to look human nature in the eye': and he sketches an ambitious programme for the future, which includes, first, the foundation of a research institute in which psycho-analysts would collaborate with clinical psychologists, neurologists, biologists, sociologists and statisticians, and, secondly a long-term research project involving the co-operation of 'a statistically adequate sample of young couples in some relatively stable community'.

In the final lecture (which might more logically, perhaps, have come first) Dr Pumpian-Mindlin, a psychoanalyst and the editor of the series, gives a lucid account of the historical origins and development of psychoanalysis and its relation to sociology and biology.

Psychoanalysis and the Unity of Science deals with the operational status of psychoanalytic concepts; Freud's theory of instinct, with special reference to the vexed question of 'alternative manifestations' (the aspect of Freudian theory which has led Eysenck to say that it can explain anything but predict nothing); the verification of psychoanalytic hypotheses; the bearing on psychoanalysis of 'perceptual rigidity' and other aspects of cognitive psychology; and the relation between psychoanalysis and ethics. As a monograph it is perhaps a trifle formless, and some of the semantic discussion is heavy going; but, like everything written by this author, it abounds in penetrating insights and constructive ideas.

MARGARET KNIGHT

The Analysis of an Obsessional. By R. W. PICKFORD. London: The Hogarth Press. 1954. Pp. xii + 234. 21*s*.

Dr Pickford is one of the very few psychologists in this country who combine distinguished work in the academic field, notably in social and experimental psychology, with a very full involvement in clinical psychology and psychotherapy. In this book he uses the detailed records of an analytic form of therapy with a patient, seen about once a week over eighty sessions, to demonstrate the clinical functions and scientific status of the method and to illustrate a good deal of its theoretical basis. It is not presented as a typical case study, but rather because of the 'interesting fantasy material produced by the patient, and the clear way in which the details of the illness and its causes and treatment revealed themselves'. A large part of the material considered consists of most interesting sequences of dreams.

This book will be criticized by members of the orthodox schools of psychoanalysis on account of the technique employed in many instances, the use of transference in the analysis, and indeed on some points of theoretical approach. The title of the book and its diagnostic label may focus some of the criticisms. But if the work reported is viewed more as an adaptation of psychoanalytic technique to the somewhat short term and non-intensive (i.e. only about once per week) possibilities in the case, Dr Pickford's method raises many fruitful issues which perhaps are of more importance than the probable points of criticism; for example, the possible advantages of shorter and less intensive work with certain kinds of patients, the need to understand far more about the dynamic structure of personality as a basis for planning treatment, and, what the author so well illustrates, the importace of clarifying why the analyst employs a particular technique in the therapy as a whole as well as in the precise situations that occur in its progress.

This reviewer finds the earlier chapters, in which more theoretical discussion is introduced, less satisfying than the more clinical descriptions. Dr Pickford's analyses of the dream sequence are particularly illuminating. The very simple style of writing belies his unusually sensitive insight into the psychodynamics of association and symbolic meanings.

In a brief review it is not possible to draw attention to more than a few of the valuable clarifications offered in this book. There are very good illustrations of the layering and interweaving of phantasies from different levels of personality development, such as are classically ascribed to the oral, the urethral and anal, and the genital phases. There is evidence of the way in which unconscious phantasy meanings influence all human endeavours, including the narrowly scientific. For the clinical psychologist there are many relevances to the use of projective methods of personality assessment; some of these are pointed out by the author, many others are clear throughout the book.

It is doubtful how successfully Dr Pickford has achieved his aim in demonstrating the clinical function of a psychoanalytic method as well as its scientific status. But only those critics from the 'scientific' opposition who have abnormally strong resistances will remain unmoved by his efforts.

H. PHILLIPSON

Dreams and Nightmares. By J. A. HADFIELD. Harmondsworth, Middlesex: Penguin Books Ltd. 1954. Pp. xii + 234. 2*s*. 6*d*.

It is recorded that two thousand years ago there were in Jerusalem twenty-four professional interpreters of dreams whose diverse interpretations of any dream were all equally correct. Dr Hadfield's skill in dream interpretation would, I feel sure, exceed their combined efforts. He gives a brief account of popular, semi-popular and unpopular explanations of dreams, physiological ('heavy supper'), physical, and premonitory theories, theories in terms of personal or racial reminiscence, and the celebrated Freudian, Jungian and Adlerian theories, and he finds a grain or more of truth in all of them. Selecting what he regards as their most reasonable features and adding something significant of his own, he advances a new explanation, which he calls the Biological Theory of Dreams. The characteristic of dreams, he holds, is their problem-solving powers. The beginning of the dream tells us the nature of the problem confronting the dreamer and perhaps its cause, and the end of the dream suggests a solution.

This theory, well presented though it is, would carry more conviction if we could assume that the dreamer usually knew the meaning of his dreams. This is indeed what Dr Hadfield claims, for he tells us that the dreamer 'can sense the significance of a dream intuitively although he may not be able to understand it cognitively' (p. 97). Among his many examples is one from a patient who dreamt that she was walking down Regent Street with no clothes on. 'By making her exhibitionistic, but telling her it does not matter, the dream quite automatically gets rid of her embarrassment and restores her self-confidence; the dream does this independently of any interpretation, merely by her undergoing the experience.'

But surely it is not enough merely to *assume*, as Dr Hadfield seems to do, that we have an intuitive understanding of our dreams. We require some evidence that this is the case.

The book is clearly and attractively written and embodies considerable experience and many shrewd observations. It will doubtless be thoroughly enjoyed by a large lay public. The critical student, however, is not likely to feel satisfied either with the chapter on nightmares or with that on dreams and psychical research. In particular he will find the discussion of neurosis and dreams inadequate for it is not based on a coherent account of infant development.

There is an error to be corrected: Heffer is the publisher not the author of *The Anatomy of Inspiration*, written by R. F. M. Harding.

JOHN COHEN

Mental Health and Human Relations in Industry. With a Foreword by Lord Horder. Edited by T. M. LING. London: H. K. Lewis and Co., Ltd. 1954. Pp. xvi + 268. 21*s*.

Of the many institutions offering training in 'human relations', the Roffey Park Institute justly holds the highest place; clinical experience, research, and a wide choice of skilled lecturers have given it an international reputation for its teaching about the psychological, social and medical problems of industry.

Now, the key points of that teaching have been gathered into the present book (it is also being published in America, and translated into ten languages from Arabic to Serbo-Croat-Slovene). Under the able editorship of Dr Ling, who contributes the introduction and four of the subsequent chapters, the book presents the substance of the teaching of some of those who have lectured at the Institute— it can be the substance only, because the wide scope of the book does not allow many of the points to be fully developed. Yet even this is largely overcome by excellent indexing, full references for each chapter, and unobtrusive cross-references between chapters; the result is a first-rate summary and source book, and a fine example of interdisciplinary teamwork, for the authors include two psychologists and two psychiatrists, two industrial medical officers, an industrial sociologist, an industrial education officer, and a former personnel manager. As Dr Ling says in his preface, 'the composition of this team symbolizes the way in which social problems in industry are being considered today'. This partnership is very happy, and an example of the value of such co-operation, with its mutual respect for different approaches to the same subject. It is not easy to single out individual chapters (and all are of a very high standard anyway), but of particular interest to psychologists will be V. W. Wilson's two chapters (on the individual and the group in industry, and on the importance of selection)—both saying much, clearly, in little; and H. G. Maule's succinct chapter on the working environment. N. C. Rimmer, in the last chapter, sums up well the main implications of the book for management.

This book, clearly written, free from jargon, readable, and above all, of topical importance, is one which should be a well-thumbed prize on the bookshelf of any industrial psychologist, or indeed, of anyone concerned with human problems in industry, which, as the book rightly says, is 'an inescapable responsibility of all grades of management'.

T. M. HIGHAM

Industrial Training for Manual Operations. By W. DOUGLAS SEYMOUR. London: Pitman, 1954. Pp. xii + 203. Price 20*s*. net.

This book is addressed to managers and supervisors and has two aims: (i) to give a comprehensible outline 'of the mental and bodily processes involved in the acquisition of skills'; (ii) to provide them with a 'guide book on how to set about training workers on repetitive operations'. The second aim is brilliantly achieved. It is not that the first aim is not achieved: it is rather that the chapter on the theoretical basis of the acquisition of skills seems irrelevant to the rest of the book. For example, the essence of Mr Seymour's contribution—or rather the basis on which his system of training is erected—is this: repetitive tasks should be trained by the 'part' method, the training being carried out at the speed of the experienced piece-rate worker; the parts are fused into pairs, then trios, then, by way of training of 'overlaps', on to the 'wholes'. As far as I can see, no light is shed on any of this by Chapter II, the theoretical chapter. In fact, the problems of wholes versus parts, or slow versus fast—these are not faced at all. It is, indeed, rather unfortunate that Mr Seymour was persuaded to insert this theoretical section: it smacks of the text in Education Psychology in which the author dutifully parades Thorndike's cats, Pavlov's dogs, Munn's rats...and Old Wolfgang Köhler and all—thinking apparently, that one must do all this before getting on with the job of showing intending teachers how to run free activity in the infant school or of telling them not to hit small children if they get their sums wrong. Mr Seymour does, indeed, warn the ordinary manager and supervisor that he can ignore this section.

The main body of the book begins with a brief, but first-rate chapter on the analysis of operator skills giving all the warnings to the amateur on the nature and amount of material required for a satisfactory analysis. The next four chapters make an excellent detailed guide to the preparation of a training course: analysis, breaking down the operation into parts, training devices for these, synthesis of the parts. There are some good tips on inspection, with emphasis given to the difficult problem of quality specification. The training devices are well illustrated with photographs which really tell you something. Then follow separate chapters on the Training Centre, its staffing, the training of instructors, organization and supervision of a training centre (an excellent record card for trainees will be worth the cost of the book to many firms) and re-training of older and 'experienced' workers.

The last chapter, called 'The Results of Operator Training and Re-Training' is a restrained claim that training pays. Mr Seymour acknowledges that most of the results he quotes are from investigations of *any* efficient training versus just-picking-it-up. One instance is, however, modestly but convincingly recorded of a factory where Mr Seymour's methods could be compared with an existing training scheme.

The 22-page appendix reproduces a detailed plan of a training course in seamless linking, again with good photographs. I shall look forward to Mr Seymour's next book, *The Nature and Acquisition of Industrial Skills*, which is now in preparation. My guess is that, given the scope of a complete book, he should be able to do something which still remains to be done: bridge the gap effectively and convincingly between the experimental psychologist and the hugger-mugger of the industrial world.

DENIS MCMAHON

The Sociology of Work. By T. Caplow. Minneapolis: Minnesota University Press; London Geoffrey Cumberlege. 1954. Pp. x+330. 40s. net.

This is an extremely good book, with virtues lacking in the bulk of American writings on social science. It is brief, direct, and has a sense of form and proportion. It is also very attractively produced; the notes are well disposed, and the bibliography first-rate, apart from the usual inadequate coverage of relevant British and European work.

Professor Caplow has been a pioneer in the teaching of occupational sociology at the University of Minnesota. In the intervals between lectures and classes, besides getting through a vast amount of reading, he has found time to give advice to large firms on their administrative problems. After reading this book, one feels sure that he gives good advice, far removed from the flabby generalizations that often pass for 'human relations'. His approach is comprehensive; embracing mobility and status, conditions affecting the choice and assignment of work, union organization, and occupational beliefs and stereotypes. (The main difference between industrial and occupational sociology seems to lie in the range of occupations covered: Caplow ranges from beauticians to bootblacks, and cartoonists to cesspool cleaners.) There is a particularly interesting chapter on the occupations of women, in which the kinds of insights familiar through the work of Mead and Riesman are supported by more references and hedged around with more qualifications than usual. Although one finds frequent anthropological and historical excursions, the major emphasis is upon American society; a formidable enough topic in itself, of course. The major omission is a study of management comparable with the treatment of labour problems and union organization. Unfortunately, when managers adopt the open-door policy and let in the social investigator they open the door to the works rather than to the back office. Readers who want to find out more about the organization and beliefs of American management must turn to such studies as William H. Whyte's *Is Anybody Listening?*, a sprightly collection of articles by the assistant managing editor of *Fortune*.

A weakness of many American works on occupational sociology, especially those in the Mayo tradition, has been the avoidance of fundamental problems of power relationships, with all their political implications. Caplow does not make this mistake; he attempts to give due weight to coercion and control in the occupational system. The theme of control, rather than coercion, runs through much of the book, possibly because most of Caplow's experience has been with large-scale organizations which have a considerable bureaucratic element.

In one chapter, an important part of the social life outside the occupational field is examined. The strains placed upon modern families by their sharp separation from the highly specialized, impersonal, rapidly changing economic order are familiar to us all, not least through personal experience. In this study of the problems involved, the material is well ordered, and a new look given to housewifery by subjecting it to a job analysis and comparison with other occupations. It is found to be a queer occupation indeed, with its own frustrating dead ends and a reward system of which the guiding principle seems to be 'the more work, the less pay'.

The publishers believe that *The Sociology of Work* will be found useful as an introductory text, in addition to providing a handy source of information for research workers. Both these claims seem reasonable. But the American text—even when it has as many passing references to other societies as this one—has obvious disadvantages when used outside America and one hopes that Professor Caplow's work will soon be complemented by detailed and comprehensive studies of the division of labour in other countries. There is no lack of raw material.

J. F. MORRIS

The Social Psychology of Industry. By J. A. C. Brown. Harmondsworth, Middlesex: Penguin Books Ltd. 1954. Pp. 307. 2*s.* 6*d.*

This readable book deals with relations within and between groups, and anybody who is interested in the human problems of industry will find it both instructive and stimulating. The emphasis is naturally on the social aspects of work and the author begins by tracing the changes that have taken place from medieval to modern times and then considers human nature and society, the formal and informal organization of industry, the nature and conditions of work, leaders and leadership, and frustration. He foresees that 'mass production techniques will be increasingly used all over the world and that the large enterprise, nationalized or otherwise, will grow in importance' and he believes the solution lies in 'the ability to adapt the new technological organization of industry so that it corresponds more closely with human needs'. According to Dr Brown, these needs are the desire for status, appreciation and emotional security; and all other factors, such as wages, welfare and working conditions, are of secondary importance. The book is intended primarily for the factory manager, the personnel manager, the time and motion engineer and the interested layman and it is commendably free from technical jargon. Some of the statements are rather too dogmatic and some of the supporting evidence is of doubtful validity; it needs time and patience to separate the substance from the shadow. Yet most people will agree with Dr Brown's diagnosis of the ills that industry is heir to though they would doubtless like to know just what should be done to cure these ills.

S. W.

The Social System. By Talcott Parsons. London: Tavistock Publications Ltd., 1952. Pp. xviii + 575. 42*s.* net.

A number of American anthropologists, psychologists and sociologists (Professor Parsons among them) met at Harvard in 1949 with the grand design of outlining the basic principles in the study of social relations. Two years later there appeared the first-fruits, the joint work entitled *Toward a General Theory of Action*; and, as its title suggests, the fundamental concept is that of action. By this is meant the behaviour of interdependent persons 'oriented to the attainment of ends in situations, by means of the normatively regulated expenditure of energy'. It follows that a general theory of action must have three components: a theory of personality focused on the individual actor, a social theory dealing with the relations of actors to one another, and a theory of culture concerned with the regulating process of cultural patterning. In *The Social System* Professor Parsons sets out to meet the second of these needs by making 'an attempt to present a logically articulated conceptual scheme' for the analysis of societies. The main body of the book falls into two parts. The first outlines and elaborates a scheme for the analysis of social structure in terms of five pairs of contrasted 'pattern variables'. The second deals with the motivational processes in social interaction and the relations of the social system to the cultural tradition. For good measure the book ends with a chapter in which the analytic scheme is applied to the American medical profession, another on social change, and one in which the spheres of interest of the various social sciences are defined.

Classification seems to be a ruling passion in Professor Parsons, but whether he will go down in history as sociology's Linnaeus is mainly a question for his fellow-sociologists to answer in due time. Certainly his taxonomic scheme will have to pass an extensive empirical testing before gaining general acceptance as the best social typology available, and before that can happen it will have to be considerably elaborated, he says. But although his suggested analysis of social structure is probably his chief contribution to the general problem, the book by no means neglects functional aspects. The debt to Freud is unmistakable here, the treatment of motivational processes being a telling adaptation of psychoanalytic theory. Implicit in it, too, is an engagingly moderate version of field theory.

This approach to the dynamic aspects of social relations should by itself entitle the book to an audience among psychologists, but two considerations threaten that audience's size. Professor Parsons freely admits that his theory is still fragmentary and rapidly developing. The temptation to wait and see what comes of it all is understandable, therefore, but it would be a pity if psychologists failed to join in this ambitious combined operation by following the lead given by Gordon Allport, Murray, Sears and Tolman, all of whom contributed to the original discussions at Harvard. The other trouble is that Professor Parsons has a prose style which reduces a reader to feelings of exasperated illiteracy. Admittedly his subject is intrinsically complex and abstruse, but his way of expounding it aggravates this inherent difficulty to the point of setting up a yawning communication gap. Happily, although much of *The Social System* is a forbidding jungle of words, its penetration has been made much easier by Professor W. J. H. Sprott, who provided an invaluable guide-book in his review article in the *British Journal of Sociology* (vol. III, no. 3. 1952). With that help even those whose interest is merely in seeing psychological theories of learning, motivation and personality in broader perspective than usual will find much that is rewarding and challenging in such chapters as those on socialization and social control. To social psychologists in particular, floundering for want of theoretical orientation, the entire grand tour could be worth the effort involved.

A. J. FLOOK

The Study of Culture at a Distance. Edited by MARGARET MEAD and RHODA MÉTRAUX. Chicago: University Press; London: Cambridge University Press. 1954. Pp. x+480. 37*s.* 6*d.*

American social research is having a tremendous boom, thanks to the Office of Naval Research and the various Foundations. Many impressive and interdisciplinary projects are in progress, and many thousands of pages of reports are being published.

This volume is a collection of papers by Mead, Métraux, Bateson, Gorer and others—the staff of a project organized by the late Ruth Benedict at Columbia, and sponsored by the O.N.R. The aim of the excercise was primarily to study the culture and national character of countries behind the 'iron curtain', with an eye to political and military applications of this information. This book reports the methods devised for the 'study of culture at a distance', and gives many examples of the kind of results obtained.

The methods reported leave much to be desired from a rigorous point of view, though no doubt they can lead to results of a preliminary and suggestive nature. The principal method is the intensive interviewing of emigrants. Despite careful internal checks on the stories told, and the attempt to obtain informants from varied backgrounds, the suspicion remains that only certain kinds of people do emigrate, and that the distorting processes of memory, sometimes after 45 years, may prejudice the results. The charge of biassed sampling applies even more to the use of projection tests, quite apart from the fact that these have never been shown to be cross-culturally valid. Another method is the analysis of literature and films. An obvious improvement on this would be to study only the most popular productions, otherwise the results may be quite unrepresentative. Furthermore even when the popular art of a society has been analysed, what does this tell us? Are slap-stick films a reaction to or a reflection of the normal life of a society?

Several of the underlying ideas of this group of writers seem somewhat unsatisfactory. For example a convincing case is not made out for the usefulness of the concept of 'national character' as applied to modern societies: no one has ever fully sampled the whole range of personalities present in one such society, or compared the different distributions in two of them. Again much reference is made to a supposed integration of 'psychoanalysis, anthropology and learning theory'. This is wishful thinking: each of the three comprises different sorts of data and different kinds of theory, and a genuine integration would entail very much more actual knowledge and a complete recasting of the three 'theories' in question.

The volume is very long, has many quotations from and references to previous works by the authors (including fifty by Margaret Mead), and contains extended excerpts from interviews and stories. This is all very interesting and suggestive, but should not be confused with the results which can be obtained by the use of those more sophisticated methods now known to social scientists. MICHAEL ARGYLE

International Review of Criminal Policy, no. 3, January 1953. New York: United Nations. ST/SOA/Ser. M/3, 28 July 1953. Pp. 162. 15*s*.

Too many United Nations documents are insufficiently noticed. This is yet another excellent product, a special issue on the proceedings of the European seminar in Brussels in 1951 on the medico-psychological and social examination of offenders.

The seminar applied itself to describing the scope and objectives of such examination. It reviewed the various types of examinations required, and it studied the question of an offender's individual rights. This question dominated the study of the problem of the communication of experts' reports to the court and the question of professional secrecy.

In nineteen technical papers the subjects dealt with include the selection of persons for examination, the legal and procedural problems, biological and physical examination, psychiatric examination, social examination and the relationship between scientific examination and treatment.

Nineteen national statements describe the existing practice under the headings of methods of scientific examination of offenders and of inquiry into their social background, the function of scientific examinations and social investigations in judicial procedure, and the treatment and rehabilitation of convicted persons.

Many of these papers and national statements, and most certainly the conclusions of the seminar, which are thoroughly comprehensive and well-balanced, should be digested by everyone working with offenders. The material studied ranges from the biological and somatic examination of offenders to the moral and psychological implications of the scientific examination of offenders. All the papers are lucid, exact and very readable. J. D. W. P.

The Young Delinquent in his Social Setting. By T. FERGUSON. Published for the Nuffield Foundation. London: Geoffrey Cumberlege, Oxford University Press. 1952. Pp. xii + 158. 10*s*. 6*d*.

The main function of this book will be its value as an antidote to loose thinking and writing on the subject of juvenile delinquency. Professor Ferguson has filled in the background of our knowledge of the young delinquent by a most thorough and penetrating investigation. He does not claim to have made any advance on, say, Burt's ideas of causation, but to have produced a balanced study of the factors influencing juvenile crime. In all 2139 boys have been studied. The results have been tabulated in 79 tables and 21 figures. This gives some indication of the scope of the work and the care with which it has been analysed.

Perhaps two conclusions might be mentioned, as they are on topics which are receiving much emotionalized attention at present. On p. 145 Professor Ferguson says: 'The fact that the mother is employed out of the home appears to exert little influence on the incidence of delinquency.' Again on p. 145 and 146 he says that among the ordinary and physically handicapped boys: '...The fact that a home was "broken" has, not, in itself, constituted a major factor in the production of delinquency.' The mentally handicapped were slightly more dependent on a full complement of parents. Again he says: 'The presence of a bad parent in the home is less to be preferred than his absence from it.'

In general the incidence of delinquency was considered in relation to a number of social variables. 'Any one of these adverse factors by itself is bad enough; but the main danger springs from a co-existence of adversities which, together, may make it difficult indeed for a lad to keep out of trouble.'

This book should prove a most useful source book for further research on juvenile crime. It is a most able illustration of the type of quantitative analysis which must be done on this type of problem before the deeper qualitative work can be attempted. MADELINE KERR

L'Enfant Instable. By ANDRÉ P. L. BELEY. Paris: Presses Universitaires de France. 1951. Pp. 43. 240 fr.

Using the method of case history, Dr Beley gives in this little book an account of eighty-four unstable boys, aged 8–20 years, for the most part delinquent, whom he saw for diagnosis and treatment at a psychiatric centre. In the Introduction he briefly discusses the concept of instability, enumerating some of the most important definitions of it. He makes it clear that he defers his own definition of the term until he has presented his data, since he wishes to do that without preconceived ideas. He is concerned with boys of eight and upwards, since instability up to the age of seven years must be regarded as normal: later than that, it is pathological. He considers that pathological instability constitutes a serious problem in modern society, since one child in four suffers from it. Such instability can be characterized by features which are either mainly physical or mainly psychological, or it may present a combination of both. It is characterized on the physical side by such signs as hypermotility, tremors, clumsy movements, tics and the like, and on the psychological side by hyperemotionality, lack of emotional development, inconstancy of thought and purpose, which may arise from conditions primarily within the child, or it may be induced by unsatisfactory conditions within the family, leading to feelings of insecurity in the child.

The eighty-four cases in his sample were studied by means of a thorough physical examination, psychological tests such as the Binet, the Rorschach, the T.A.T., level of aspiration and attention, and by means of a psychiatric examination. Their general behaviour was also under close observation, and a full family history was taken. On the basis of the data so gathered, Dr Beley was able to divide his cases into the following broad types: (i) instability of physical origin and manifesting itself in physical symptoms; (ii) instability where character involvement is present. This second group is by far the largest, fifty-six out of his eighty-four boys falling into this broad category, which is further subdivided into ten groups. The principle of classification adopted by Dr Beley is not always very clear, and it is difficult to gather from the illustrative data presented why Dr Beley should have included in this group such heterogeneous cases as the hyperkinetic, the mentally backward, the schizoid, the paranoid, and the epileptoid; (iii) instability where there appears to be no connexion between the physical status of the child and any disorders of behaviour, and where the origin seems to lie in family conditions.

Dr Beley has undoubtedly been able to gather together a great deal of information of interest, but a different presentation of results would have allowed the ordinary reader to get a clearer picture of the conditions he describes. He gives short case histories of thirty-eight of the cases selected as being the most representative, but gives no tabulated information about the remaining cases. For instance, their ages are not given. Only six of the thirty-eight cases cited are below the age of twelve years; and it is clear that much of the investigation was concerned with the problem of the unstable adolescent rather than the unstable child.

Although a number of psychological tests appear to have been used, only the results of the Binet are given (and not in every case), and then only in terms of mental age. Scatter is mentioned as characteristic of the majority of cases, but no details are given. Further, from examination of the mental ages supplied, it seems clear that the majority of cases are of poor intelligence. Whether this is true of the sample as a whole it is impossible to say.

The other data presented are given in an inconsistent manner. Family heredity and details of family background are given in some cases and not in others. This makes comparison of the cases cited difficult, and one is given no help, since even the simplest quantitative treatment is omitted. Although the sample itself is quite large, it will scarcely stand Dr Beley's technique of subdivision. For example, in some instances only one or two cases fall into each subdivision of his second broad class, and no deductions have been or can be made as to their significance.

There are a few words of summary at the end of each section dealing with particular types of instability, but the general summary and conclusions at the end of the book are vague, and there is little attempt to discuss the concept of instability in the light of his own investigations.

It seems clear that considerations of space must have forced Dr Beley to be unduly selective in his presentation of the material he has gathered. This is to be regretted since his careful methods of observation and testing must have yielded a great body of interesting data.

GERTRUDE KEIR

The Psychopathic Delinquent and Criminal. By GEORGE N. THOMPSON, M.D. Illinois: Charles C. Thomas; Oxford: Blackwell Scientific Publications. 1953. Pp. xv + 157. 32*s*. 6*d*.

The study of the psychopathic delinquent comprises the most valuable research in criminological psychiatry as, if for no other reason, it is this individual who has repeatedly precipitated nations into war.

The author believes that there is more to human behaviour than simply the working of a series of chained conditioned reflexes. He holds that a function of the diencephalon is the selection of response. This book, which studies 'a non-psychogenic disorder at a time that interest in psychogenesis is at its peak', should be read by psychologists as well as psychiatrists, who have been attempting to apply psychogenesis to almost everything. The author's argument is well disciplined.

The specific defect of the psychopathic personality is held to be the inability of the person to develop an adequate concept of time in relation to self. All the secondary symptoms stem from this. It is his disregard for the consequences of action that make the person with a psychopathic personality such a puzzle.

The author adduces much evidence to show that psychological factors seem to have little to do with the aetiology of the true psychopath. The statement by Nielsen and Thompson that 'the most important function of the environment...is to inhibit, or augment psychopathic behaviour by the opportunities it affords' stands essentially unchallenged, but environmental psychogenic factors lead to delinquency both in psychopaths and non-psychopaths.

An excellent picture of the characteristics of the psychopathic personality is given in a separate chapter and four main types are described—pathological emotionality, asocial or amoral trends, pathological sexuality, mixed types—with many subdivisions which are not too convincing.

In discussing treatment, a plea is made for separate units or special hospitals for these patients. Psychotherapy benefits only some cases in which the concept of 'partial psychopathy' is applicable. Chemotherapy is frequently useful when there are cortical lesions. The author thinks that perhaps the most hopeful therapy will be found in the field of neurosurgery.

Professor Thompson has made an important contribution to the literature and this book should be studied not only by those who are organically orientated, but more particularly by the psychogeneticists.

J.D.W.P.

Maternal Care and Mental Health. By J. Bowlby. Geneva: World Health Organization. 1951. Pp. 179. 10*s.* net.

It is too late to recommend this book to readers for it has recommended itself, and together with the shortened Penguin edition has reached the dimensions of a 'best seller'. The interesting question which arises from this fact is how to account for its popularity and the influence which it has undoubtedly had on public policy.

Dr Bowlby's theme is a simple one. It is that harmful and even dangerous effects can be the result of the separation of small children from their parents. In support of this idea Dr Bowlby collects and documents the work of clinicians and research workers throughout the world.

In this country *Maternal Care and Mental Health* has been widely read by an enlightened public which was shocked at the clearly presented facts about the effects of deprivation. This was the same public which had previously been shocked by the findings of the Curtis Committee, and it seems certain that much of the interest in Dr Bowlby's book is derived from the feelings aroused by the inquiries of that committee.

When *Maternal Care and Mental Health* was published in 1951 enough time had elapsed since the passing of the Children's Act in 1948 for the problems of caring for children deprived of a normal home life to have become evident to all those concerned with the work. The first enthusiasm which had gone into the organization of the Children's Departments within the Local Authorities had worn off as the difficult nature of the task revealed itself. It was just at this moment, when the need for a clearer understanding of the problems of deprived children was keenly felt, that Dr Bowlby's book appeared. It was seized upon and read with avidity by thousands of committee members, administrators and professional workers of all kinds, and there can be no doubt that a new awareness of the emotional needs of children was gained by many people. All those especially concerned with the work of caring for deprived children must be grateful to Dr Bowlby for his timely presentation of knowledge which was urgently needed.

But not only did Dr Bowlby give people the knowledge which they were seeking, he also gave them the feeling that there was something which they could do about the suffering caused by maternal deprivation: they could take more active measures to prevent the separation of children from their parents. The present emphasis on preventive aspects of child care owes much to Dr Bowlby's influence.

Because the fact presented by Dr Bowlby came as a shock to many people, there was a reaction to them, as there always is to shock. Reaction takes the form of either denying the facts, or accepting them to the exclusion of other facts. And reactions of both these kinds are among the snags about a book like Dr Bowlby's, for present policy can be seen to be influenced by people who have reacted in one or other of these ways.

The difficulty is that people who merely react to new facts tend to use them as a basis for argument one way or the other, and in order to gain political ends if necessary. They fail to assimilate the new ideas and so do not gain the further insight which leads to new development. It is this element of shock in Dr Bowlby's book which could prevent the assimilation of the truths it contains, and could hinder new growth in understanding.

CLARE BRITTON

International Review of Criminal Policy, no. 4, July 1953. New York: United Nations. ST/SOA/Ser. M/4, 19 October 1953. Pp. 110. 15*s*.

The first two articles deal with sex offenders. G. K. Stürup discusses their treatment in Denmark and the other Scandinavian countries. In particular he discusses the effects of therapeutic castration and concludes that they are exceptionally good. Karl M. Bowman reviews sex legislation and the control of sex offenders in the United States of America. He concludes that it seems better to reform procedural rather than substantive law, premature change in which is always dangerous. In one experimental treatment service psychiatric treatment was supplemented by religious guidance.

Three papers on the treatment of adult offenders in Australia, Venezuela and Yugoslavia are of little or no interest to the psychiatrist or psychologist. There is, however, a comprehensive topical bibliography of current periodical and non-periodical technical literature, concerned with the prevention of crime and the treatment of offenders, comprising 1227 publications.

J.D.W.P.

The Modification of International Attitudes. A New Zealand Study. By J. R. McCreary. Victoria University College: Publications in Psychology No. 2. 1952. Pp. iii + 153.

This monograph reports a study carried out in New Zealand in 1950 as part of the U.N.E.S.C.O. project 'Tensions Affecting International Understanding'. The aim was to test the effect of knowing more about the Japanese upon attitudes towards them. A Bogardus-type Social Distance Scale was used before and after information had been given to the subjects. One thousand subjects, both adult and child, took part in the first phase of the study, but wastage was high, and only about half of these completed all three phases. The sample can in no way be considered representative of the New Zealand population.

The conditions under which information about aliens was found to affect attitudes include active participation in obtaining such information; the film rather than lecture or booklet mode of presentation; in the case of presentation by speakers, intimate knowledge of the alien group, especially as a member of it, together with equal or superior social status to that of his hearers. The result of presenting 'all sides' of the question was equivocal, and the effect of University classes negligible. Where a national group is the centre of well-established antipathies, specific rather than general information seems to be the more effective.

As the investigators point out, the study is weakened by lack of control groups, and by the short period over which change was measured. No protocols are included to show in detail how the subjects were instructed, and so one can make no guess as to how far the idea of participating in the experiment may have swayed the subjects' opinions. Otherwise, the report is very clear and well organized, and the discussion of remaining problems promises well for future studies from Wellington.

CECILY DE MONCHAUX

The Makah Indians. By Elizabeth Colson. Manchester: University Press. 1953. Pp. xvi + 308. 28*s*.

This excellent book is a study of the degree to which the Makah Indians, a population of some 400 who live in a small reservation on the coast of Washington State, have been assimilated to American culture. Its essential theme is, therefore, not the process by which they have come to lose most of their characteristic customs, but the extent to which they remain and consider themselves a distinct group. They are no longer distinguishable by language, physical type or mode of life from other Americans living among them, and these other Americans must be regarded as forming part of one community with them. Yet they are intensely conscious of their separate identity. The basis of this common consciousness in Dr Colson's view is their identification with the reservation, an area in which they have certain vested interests not shared by white residents. It is maintained by an ideology which represents all whites as actual or potential despoilers of all Indians and as prejudiced against them; the Indian customs which the officials of the United States Government did deliberately suppress during a period of some seventy years are included among the goods of which they have been despoiled, and although no one now remembers the days when they were practised, the knowledge of their having existed is an important part of the cultural heritage. From her own observations Dr Colson concludes that although there is some social segregation of whites and Indians in the reservation, there is no general prejudice against Indians as a group, such as exists outside and makes it difficult for them to rise above the lowest economic classes. Her study is a valuable contribution to the literature on the position of minority groups; it also suggests depressing parallels in the attitudes of peoples in other parts of the world to whom Western civilization has been brought as an element in alien rule.

L. P. MAIR

Psychological Testing. By ANNE ANASTASI. New York: The Macmillan Co. 1954. Pp. xiv+682. $6.75.

The book is divided into four parts. Part I, which occupies 170 pages, deals with the principles of psychological testing. This includes the nature and interpretation of norms, reliability, validity and item analysis. These principles are then applied carefully and clearly to the various tests discussed in the remaining three parts, which are concerned with General Classification Tests, The Differential Testing of Abilities and the Measurement of Personality Characteristics. These sections are given 173, 175 and 136 pages respectively. This perhaps affords some indication of the author's position in relation to Factor Theories. A rough assessment of her general attitude might be given by saying that the evaluations of the many tests would probably yield a pessimistically skewed curve on a three-point scale of Preposterous, Poor and Promising.

The book is written in a style very much more familiar to British readers than is customary in American publications on Psychology. Such readers may, however, be somewhat disturbed at finding no reference to Burt, Thomson or Vernon in the discussion on Factor Analysis and only about three or four references to British tests. Readers in other European countries may be similarly dismayed. If the Monroe Doctrine is to apply to Psychology, Professor Anastasi's book would lose nothing by being re-named *Psychological Testing in America*. In spite of this ruffling of European pride, it is unlikely that serious students of psychological testing will find this book other than skilful in presentation, balanced in judgement and generally much too good to miss.

G. A. FOULDS

The Appraisal of Intelligence. A. W. HEIM. London: Methuen and Co. 1954. Pp. vii+171. 12*s.* 6*d.*

The first page of Miss Heim's book explains that 'appraisal' may be contrasted with *measurement* on the one hand and with *subjective judgement* on the other. In this way the keynote for the whole book is struck firmly at the outset, and the reader is prepared for a sustained attack on the psychometrist, the factor analyst and the group tester. The main thesis of the book is that 'intelligent behaviour may admit of differences in kind as well as in degree', intelligent behaviour consisting of 'grasping the essentials in a given situation and responding appropriately to them'.

Readers of this Journal are well aware of the differences in opinion and attitude between the clinical psychologist and the group tester, between the pure psychologist and the applied psychologist and between the factor analyst and the psychophysical laboratory worker. Each of these contrasts is capable of producing lively controversy: when three of these roles are combined in one person the resulting attack is likely to be pretty strong, and mounted on a wide front. The sins of the psychometrist, the applied psychologist and the factor analyst are described with exhaustive fluency and no quarter is given. The only psychologist to emerge from the fray with credit is Binet—indeed, the author appears at times to lean over backwards in her desire to give full credit to his genius. It is inevitable that in the heady intoxication of combat Miss Heim is occasionally led to the setting up of papier-maché ogres just for the joy of laying them low, and that in a number of places the writing becomes tendentious and perhaps a little querulous.

Yet this is a book which was well worth writing, and one which is strongly recommended for students as well as psychometrists. There are not enough books of this kind available: attacks on intelligence tests and testers are too often written by those who do not know their subject and who are incapable of appreciating the technicalities behind the methods they attack. The chapters dealing with the nature of intelligence and the concepts of mental age and I.Q. will be of considerable value to students, as will those covering reliability and unreliability. These chapters, perhaps, could be expanded a little in any future edition—the consideration of the standard error of score, for example, would be a useful addition. As one might expect from Miss Heim's own research work, the section on speed versus power is of considerable interest. Much of what Miss Heim has to say about validation and its difficulties is excellent, but one feels that the omission of any reference to follow-up studies as a validating technique makes the discussion less valuable. For the educational psychologist, the 'validity' of an intelligence test is often to be found in its predictive power. Perhaps Miss Heim would find this an inadmissible use of 'validity'—if so, she ought to argue the case.

The final two chapters make more constructive suggestions after the largely destructive criticism of the previous eleven, and here many pertinent and suggestive points are made for future research. It is interesting to note that many of these points—e.g. difficulty of items, speed, the use of creative instead of choice responses—are those which in the educational sphere have recently begun to receive more attention. Miss Heim's suggestion for new methods of appraisal of intelligence will appeal particularly to the clinically-orientated psychologist working with adults. Most of these suggestions need working out more fully, with a particular eye on a number of the criticisms made by Miss Heim herself in earlier chapters. May we hope that these suggestions are an indication of future lines of research by the author?

STEPHEN WISEMAN

Studies of Psychomotor Personality Tests I. By MARTTI TAKALA. Helsinki: Institute of Occupational Health: Ser. B. 81, No. 2. Annales Academiae Scientiarum Fennica. 1953. Pp. 130. 400 marks.

The work carried out at the Institute of Occupational Health, Helsinki, which is reported in this monograph, falls into four parts. The first, concerned with the establishment of factors of motor performance common to different tests, reports results obtained from the factor analysis of variables derived from psychomotor tests given to a group of student subjects. Of the five factors resulting from the analysis, three seem to be of importance. These are identifiable as being concerned with (1) speed, (2) smoothness *v.* disturbance, and (3) suddenness *v.* sluggishness, of psychomotor reaction. The second factor is thought to be related to a general neuroticism factor, whilst the third, it is tentatively suggested, may be analogous to an overactivity-inertia type factor. The second part of the work is concerned with a very thorough analysis of the Mira test, the results of which should be of value to anyone intending to employ this technique. An attempt is made in part three to relate Mira test variables to the three factors outlined in the first part. Results indicated that some wider interpretation of these factors may be necessary. The relation of the psychomotor to other personality tests is treated in part four. This yields some additional information, but on the whole correlation coefficients are not higher than usual in this type of work; a state of affairs which the author rightly says should lead to further development of research methods with attempts to improve the reliability and validity of the tests. The assessment of the work reported is made more difficult than necessary by the use of non-standard scoring methods for tests for which data are available in this country. It is also to be regretted that the work was not carried out within the framework of some general hypotheses, which would have permitted a clearer evaluation of this otherwise careful work.

PETER VENABLES

Growth and Development. By KARL C. GARRISON. London, New York and Toronto: Longmans, Green and Co. 1952. Pp. xii+559. 32*s.* 6*d.* net.

This book gives an account of every aspect of child development, physical, intellectual, emotional and social, and provides a good reference book on the subject. Much modern work is described, and the bibliography is most valuable. The form of the book is not perhaps very congenial to British readers, with its list of questions for the students at the end of each chapter. It gives too, a discrete picture of the child's development, and there is not enough emphasis on the interrelatedness of the different functions, nor on the dynamics of their growth and change. Its greatest value would be to the teacher or other child worker in training, as a source book.

A.M.S.

The Child's Conception of Number. By JEAN PIAGET, translated by C. GATTENGO and F. M. HODGSON. London: Routledge and Kegan Paul. 1952. Pp. ix+248. 25*s.*

This translation, first published in 1952, keeps very close to the original *La Genèse du Nombre chez l'Enfant* and, at the same time, describes the experiments in terms which read quite naturally to an English reader. This is no mean achievement since the environment of the children and, consequently, the small details of the experiments and the idioms in the answers might have seemed somewhat unreal had it been less well interpreted. This seems to have been accomplished with no modification of meaning which I could detect.

The book deals with the development of fundamental bases of arithmetic in children up to the age of about seven years. The stages by which an idea of conservation of quantities is built up are explored in great detail and anyone concerned with the teaching of young children would be well advised to know what is in the first two chapters before attempting to introduce any 'units' into the arithmetic lessons. Four chapters are devoted to probing the growth of an understanding of cardinal and ordinal numbers, in the course of which a most important analysis of the concept of 'seriation' sheds much light on what is involved in a simple ordering of a set of quantities and in establishing a one to one correspondence between two such sets and in the relation between these two processes. The final section is devoted to relations between numbers, but, characteristically, Piaget going more deeply than the simple number combinations, probes into the concepts which must be developed before these combinations can be meaningful. His chapter on the relation between parts and wholes is particularly informative.

This is a thorough and penetrating study of the concepts which are being formed in relation to numbers and quantities at what may be called the 'Infant School' age and, as such, it throws considerable light on the difficulties encountered at this stage when the foundations of future work in arithmetic and mathematics are being laid. The book, however, has a significance far greater than this because it is the first (published in French in 1941) of a series of books devoted to a long and thorough study of the development of arithmetical, spatial and temporal concepts including such books as *La représentation de l'Espace chez l'Enfant* and *La Géometrie Spontanée*. In his earlier works, beginning with *The Language and Thought of a Child* in 1924, Piaget gathered together an enormous quantity of what one may call clinical material, concerning the development of various concepts, beliefs and techniques in children. Piaget was interested in the development of individual children and spurned the 'mean-child' and the standard deviation, making no use of the familiar techniques of statistics to bring one sort of rigour to his researches. The importance of *The Child's Conception of Number* may well lie in the fact that it is the first of a long series of books in which the psychological findings are fitted into a system based on the rigours of mathematical logic. The series culminates in *L'Épistémologie Génétique* (1950) in which he discusses at length the theories of number and spatial relationships, the philosophical basis of his work and the psychological processes involved in an understanding of logical classes and 'groupements', a word which Cattengo and Hodgson translate as 'groupings', but which is used in a special sense and might well have been left untranslated. Considered in relation to the rest of his work, this book marks the beginning of an important stage in Piaget's work and possibly of a fresh advance in psychological theory, particularly in its relation to philosophy.

F. W. LAND

Logic and Psychology. By JEAN PIAGET. Manchester University Press. 1953. Pp. xx + 48. 6*s*. net.

In this little book, based on three lectures given in Manchester in 1952, Piaget indicates, in a thumbnail sketch, the very important and significant light which, in his view, modern logic can throw on genetic psychology. His position briefly is that modern operational logic can supply just the required explanatory theory for all the observed facts (as established and recorded by him and his co-workers over the last thirty years) of the stages, the diverse aspects and the final outcome of the great, long-drawn-out process of intellectual development in the child. In particular the 'algebra of logic'—translated into a suitable intermediate 'psycho-logic' or 'logico-psychology'—can make intelligible to us why various *seemingly unrelated* types of achievement present themselves *at the same time*, at the age of about twelve: in the light of this algebra of logic they now all fall into place and we can see them as the naturally correlated members of a single whole structure or system.

The master-key for Piaget lies of course in the notion of operations. This is for him common both to psychology and to logic and the crucial link between them. After a brief account of the 'History and Status of the Problem', he outlines in his first main psychological chapter, how up to the age of about seven years the ground is prepared; how from 7 to 11 years operations of 'combining, dissociating, ordering and the setting up of correspondences' arise and develop, but are only applied concretely to objects themselves; but how from 11–12 to 14–15 years these operations become internalized, are applied to propositions instead of physical objects and above all achieve the status of 'structured wholes' within which the all-important property or law of *reversibility* holds. At the same time, however, there appear the above-mentioned new groups of operations of 'operational schemata' without any obvious relation to the 'logic of propositions': combinations and permutations, proportions, conceptions of mechanical equilibrium, probabilities, etc. These present us with an intriguing psychological problem, but here Piaget finds the illuminating solution in the 'algebra of logic' which enables us to understand the nature of the system of logical equilibrium towards which our whole intellectual evolution must tend. Accordingly he goes on, in his next chapter, to give a summary (and necessarily technical) sketch of the algebra of logic, and in the final chapter shows how this serves to explain first, the 'outline structures' of the 7–8 year level, and secondly, the full structure of *reversible transformations* of every kind governed by the laws of equilibrium and representing the attainment of mature adult logic.

At various points in the exposition controversial questions arise, but within the compass of three lectures Piaget could hardly aim beyond a simple introductory presentation of his point of view. Its solid foundation must depend, on the one hand, upon the vast body of his practical work, and on the other, upon his comprehensive theoretical treaties on logic and on genetic epistemology. How conclusively these establish every part of his position falls outside the scope of this brief notice. But it may be hoped that the present thought-provoking introduction will, for many students, serve as at least a gateway into a great world of cumulative psychological exploration and understanding, the significance of which has not yet, in the writer's view, been properly appreciated. Piaget's thirty years of systematic investigation and mapping of the developmental history of the child's mind, in every one of its intellectual aspects, will, I believe, be acknowledged in future days, if not at the present time, as far and away the most important body of contributions to cognitive psychology within our generation. NATHAN ISAACS

Statistical Methods in Experimentation. By O. L. LACEY. New York and London: Macmillan and Co. 1953. Pp. xiv+249. $4.50.

A Statistical Primer. By F. N. DAVID. London: Charles Griffin and Co. 1953. Pp. xi+226. 22*s.* net.

Both these books are text-books for subsidiary courses in statistics. Prof. Lacey's (University of Alabama) is intended for use in a department of psychology; Dr David's (University College, London) is provided for a more miscellaneous audience of 'post-graduates and undergraduates drawn from nearly all science departments and possessing little in common except an ignorance of mathematics'. Actually both assume their students have reached about the same level, algebra up to General Schools Certificate standard approximately, though Prof. Lacey evidently hopes a few of his students may have been introduced to integral calculus.

Both describe properties of the normal univariate frequency distribution and methods of testing the significance of differences between samples. Dr David goes on to consider binomial and Poisson distributions in some detail and concludes with uses of the chi-squared test. Prof. Lacey deals much more briefly with enumeration statistics and devotes his later chapters to properties of bivariate frequency distributions—correlations, regression coefficients, etc.

Perhaps it is impossible to convey complete conviction while rigorously avoiding mathematical proof. Dr David adopts an impressively authoritative manner, Prof. Lacey is more sweetly reasonable and persuasive, but neither succeeds completely. For instance Dr David, on the comparative merits of means, medians and modes, tolerates little argument: 'The first measure we shall discuss and the one we shall afterwards always use is the mean.... The real reasons why statisticians do not favour (the median) are that it is mathematically intractable and tends to be more unreliable than the mean.... The same objections can be made against the mode.' In contrast, Prof. Lacey's presentation of the formula for product moment correlation (abbreviated here) gently leads the reader along: 'A measure of the degree to which values above and below the average are associated is immediately available by the simple device of multiplying together paired deviations.... The simple sum of the cross products will be an index of the degree of correlation.... It is plausible then that the first step is the calculation of Σxy. Now it is clear that this sum is not immediately the correlation coefficient itself. It must be compared with some other figure also dependent upon the data. It may occur to you that it could be compared with either Σx^2 or Σy^2.... The value which would seem reasonable is the average of the two...it is, however, the geometric rather than the arithmetic average.... So our correlation coefficient is finally defined', etc. In both contexts an attempt to define the criteria of the efficient statistic might have been made. Without it, statisticians' preferences must appear rather arbitrary to the student.

A great variety of problems have to be met in preparing lectures for such subsidiary courses. Though there is some agreement about the subject-matter to be included, no plan is likely to prove equally satisfactory in all localities. To hold his students' interest the lecturer will probably prefer to borrow his examples from experiments conducted concurrently in his own department, and adapt his computing methods to the apparatus available there, integrating his course with the rest of the work in every way he can. So though other lecturers may find it well worth while to study how these authors present their subject-matter, I cannot estimate how widely either book will be adopted as a text-book. Students who wish to brush up statistics for an examination may perhaps find Dr David's severely practical text more suitable for their needs; those who are not yet sure how far statistics can help them in their other studies may find Prof. Lacey's the more encouraging. PATRICK SLATER

Psychical Research Today. By D. J. WEST. London: Gerald Duckworth and Co., Ltd. Pp. 144. 1954. 12*s*. 6*d*.

This little book is an attempt to describe briefly the present state of psychical research and the main problems which are now engaging the attention of students of the subject. Dr West is commendably cautious in his approach and the beginner will find in his work a useful text-book from which to start a further inquiry. In nine concise chapters, supplemented with useful and well chosen documentation, the author sketches the principal phenomena and subjects the evidence to critical analysis.

In his treatment of spontaneous psychic impressions Dr West shows himself keenly alive to the weakness of much of the evidence, although it is somewhat surprising that he omits to mention the early criticism of Gurney and his associates in their records of alleged veridical phantasms. In his account of the phenomena of the seance rooom the author exhibits a scepticism much greater than that he displays in his treatment of extra-sensory perception. It is difficult here to contradict anything that he writes: the melancholy story of fraud by mediums and suppression of the facts by psychical researchers is only too well known. Perhaps a wider acquaintance with the physical phenomena might have led him to modify some of his conclusions, or at least to have omitted one or two pages and inserted in their place a discussion of the best evidence.

It is in his treatment of extra-sensory perception that the author exhibits his keenest appreciation of modern tendencies. He is of the opinion that 'almost everyone is agreed' that E.S.P. has been shown to exist and in dealing with the evidence for it he shows himself well acquainted with the literature in English and the experimental procedure. Generally speaking, his treatment is restrained, although it is curious, if not a little disturbing to find that among the tests he thought worth recording was the American experiment conducted by Professor Riess. After describing the test and the amazing accuracy of the subject's guesses, he concludes by saying that 'it is perhaps unwise to pin too much faith upon it'. It is indeed.

The book is well printed and indexed but is somewhat marred by the illustrations which it is difficult to suppose could have been chosen by Dr West and should not have been sanctioned by him. Nevertheless, the book is to be commended to all interested in these matters and students should note one of the author's final conclusions that 'the whole subject is one big question mark'. Few, certainly, could quarrel with that. E. J. DINGWALL

The Doors of Perception. By ALDOUS HUXLEY. London: Chatto and Windus. 1954. Pp. 64. 6*s*.

Most psychologists have probably at some time or another become interested in mescalin, either by experiment and experience, or at least by some study of the large literature which has accumulated concerning results following its administration. Also most psychologists will be familiar with the claims that have been made that substances chemically closely related to mescalin can produce effects similar to, or identical with, some of the leading symptoms of schizophrenia. It is important to remember that Mr Huxley was apparently well acquainted with much of this earlier work, and that in particular he was already deeply interested in the adenochrome experiments before he tried to discover something about the effects of mescalin upon himself. What he did was to take four-tenths of a gramme of mescalin dissolved in half a glass of water, and then to record observed results in detail, using verbatim automatic recording whenever possible.

The essay, deeply interesting to read, is literary in manner, with here and there bits in the scientific style which do not altogether fit in well with the rest. It is not at all easy to tell how much of what Mr Huxley has written is due to reflexion upon his actual first-hand experience, or to the literary and stylistic demands themselves, and how much of it is really literal report. He claims that an hour and a half after he took mescalin a lot of quite ordinary things—flowers in a vase, books on his study shelves, a plain typing table, folds in his trousers, were seen as possessing some vast and tremendous, though not readily definable, significance. To Mr Huxley, then, and apparently afterwards as well, this seemed to be 'better' perception, revealing something which is known to be real in some intimate and unquestioned sense, for the description of which the perceiver must draw upon the language of mysticism.

Later, Mr Huxley moved about, in a garden and in the larger outside world. Neither space nor time seemed very important, apparently because 'there is such a lot of them'. The distinction between the world outside and the private world still remained, but the sort of world Mr Huxley got when he shut his eyes—'brightly coloured, constantly changing structures that seemed to be made of plastic or enamelled tin'—appeared to him to be cheap and trivial.

This is one of the places at which the essay is not very clear. A little later the schizophrenia theme comes in. And then it seems as if an outward directed perception—in this case of a garden chair, for example—may yield something that is ineffable, supremely important, and neither 'inward' nor 'outward'; for there simply is nothing else. Initially, it seems to be held, this is the heavenly phase of schizophrenia, but it may go on to be compulsive, obsessive and may then turn into the 'hellish' phase. But Mr Huxley is satisfied that mescalin does not take anybody that far unless he has had a recent attack of jaundice.

The final part of the essay is concerned to discuss the ways in which mescalin and similar drugs might be made a valuable adjunct of religious ceremonial. It might 'shake people out of the ruts of ordinary perception' and so make more accessible a world of values far removed from the conventions of common sense. It seems to be implied that, whether by the help of mescalin or some other drug, whether by the 'sunset touch', or by some special sensitivity to sound or shape or colour unmediated by anything else, everybody ought to get into this world from time to time and nobody ought to stay in it very long.

In a technical, scientific way there is very little psychology in this book, and 'perception' in particular is being used in the special sense of some function which is, so to speak, capable of absorbing, and being at one with, whatever, at other times, it would be said to be directed upon. But psychologists should read this essay, not merely because most of it is beautiful writing, but also because it raises again those old issues of how, if it can be done, the values of those formulations which pass as the everyday currency of meaning from person to person, may be reconciled with the other values of becoming-aware-of meanings that are not, and perhaps cannot be, formulated at all. The issues are raised, but they are not revolved.

F.C.B.

A Survey of Public Opinion in Relation to the University. By MARJORIE N. DONALD. Publications in Psychology No. 4. Department of Psychology, Victoria University College, Wellington, New Zealand. 1953. Pp. 90+lviii, Appendices. No price given.

The survey in this booklet was the practical outcome of an undergraduate course on the design and administration of public opinion surveys. The students participated at all stages, from the construction of the interview schedule to the analysis of results.

The aim of the research was to uncover the views held by a stratified sample of 400 people in the capital city of Wellington, on the subject of the University and its function within the community. Among the topics investigated were the following: Who should enter the University? What value is placed upon a University education? How does the behaviour of University students compare with that of other young people of the same age?

Detailed breakdowns of opinion in all areas are clearly presented and indicate that the local college is held in high esteem by the people of Wellington. Among the many interesting findings were two strongly held views, 'that a university training equips a student to become a better citizen than he otherwise would be' and 'that the value of a girl's university training remains with her after she marries.'

ALAN RICHARDSON

Authority and Freedom. By R. H. THOULESS. London: Hodder and Stoughton. 1954. Pp. 124. 8*s.* 6*d.*

This, addressed to theologians, is intended to be a contribution to reversing the 'decline in religious conviction'. It gives brisk support to the Creeds and to the idea of dogma but recommends the (Anglican) Church to treat most of its doctrines as 'orthodox but non-essential beliefs'. Religious intolerance is deplored in a chapter which also gives some very fine arguments for it.

The first chapter says that there has been a decline in church membership, excepting that of the Roman Church; that figures of such membership are a good measure of religious conviction; that there are indications that the decline has been somewhat reversed since 1935. It suggests that the social consequences of this decline may be undesirable; says that it is not known that delinquency is increasing but that there is some reason (not given) for thinking it may be and that the decline is a cause of the increase; it says that religion should be taught not as a means of good citizenship but because it is true. There is argument here of the sort that was roundly condemned by Dr Thouless in 1939 as 'fantastic anecdotalism'. He does not wonder *why* there has been no decline in Roman Catholicism nor what is the incidence of maladjustment among Catholics.

The last chapter attempts to resurrect the traditional arguments for the existence of God as *inductive* arguments having a probability, a howler of a 'category-mistake'.

Liberal-minded churchmen will like this book: other Christians may derive comfort but no help in recruitment, for it is not that unbelievers mistake the meaning of 'will' in 'Whosoever will be saved...' but that they have no idea what 'be saved' means, and this is *not* because they have lost the ability to understand metaphor (cf. chapter 5). Dr Thouless is described on the dust-cover as Reader in Educational Theology.

JOHN SYLVESTER

Jung's Psychology and its Social Meaning. By IRA PROGOFF. London: Routledge and Kegan Paul. 1953. Pp. xx+299. 18*s*.

The aim of this book is stated as the interpretation of Jung's ideas in such a way that they can be 'analysed, mulled over, criticized, appreciated, and most important, *worked with* as hypotheses in related fields of psychological and *social* study.' The exposition has been handled ably, as anyone who has struggled with the works of Jung would wish to testify. On the critical side some disappointment may be felt. For instance, the author shows that Jung has been more concerned with the social foundations of man's life than is commonly supposed, but he does not consider the influence of nineteenth-century thought in producing a confusion between social and biological inheritance in Jung's distinctive concept of the Collective Unconscious; on the contrary, the language used suggests an acceptance of the Jungian position. In general, the inconsistencies of Jung's writings are acknowledged more openly than is usual with his sympathizers and supporters, but their customary standpoint is maintained, i.e. that the value of his work lies mainly in 'an attitude toward the reality of the psychic life and the validity of the spiritual experiences of man in history'; and that the material which he studies needs to be experienced not only by the intellect but 'more comprehensively, by the personality as a whole'. In view of the aim as declared, more might have been done to determine, point by point, what the set of values really is which Jung is endeavouring to convey, and to explain in detail how his concepts are to be used as hypotheses in the study of specific problems. One of the most interesting features of the present work is a chapter on 'The Historical Situation', based partly on conversations with Jung himself, in which the development of his ideas and his relationship with Freud are treated more fully than in other published sources.

A. M. DRY

Steps in Psychotherapy. By JOHN DOLLARD, FRANK AULD, Jr., and ALICE MARSDEN WHITE. New York: The Macmillan Company. 1953. Pp. ix+222. $3.50.

This book is disappointing. It achieves its stated aims—the presentation of seventeen hours of therapy on Mrs B. whose fundamental 'sex-fear conflict' manifests itself as a 'spouse-phobia'. Transcripts of the sessions are given, with a commentary from an arm-chair supervisor. In the Brief Psychotherapy described, no fundamental change in psychotherapeutic technique is presented but 'merely a limited attack on the total psychotherapeutic problem'...(which)...'must necessarily deal with everyday things'. Thus, to its brevity can be added the characteristic, 'superficiality', and we have the conventional supportive therapy practised in most of our busy psychiatric outpatient departments.

The disappointment comes from the fact that while the authors state that their therapeutic techniques are based on the Dollard and Miller blend of reinforcement learning theory and psychoanalysis, the comments and records of the sessions show only a little Freud, less Hull, and no Dollard and Miller.

The appearance in 1950 of *Personality and Psychotherapy* by Dollard and Miller was hailed as a 'marriage of psychoanalysis and learning theory'. The technique revealed in this book does little to reassure us that the union is potentially fruitful, or even happy.

The final two chapters discuss the relevance of the psychological tests given to Mrs B. The reports of the clinical psychologists are equivocal. The authors feel that this is because of the lack of validity studies of such tests 'utilizing criteria pertinent to personality diagnosis and therapy'. This is all too true. But the handling of Mrs B.'s case does not demonstrate how psycho-analysis strengthened by the rigours of learning theory can provide any such adequate criterion.

R. M. MOWBRAY

Psycho-analysis and Child Psychiatry. By EDWARD GLOVER. London: Imago Publishing Co. 1953. Pp. 42. 6*s*. net.

This is an expansion of an article previously published in *Samiksa*, the journal of the Indian Psychoanalytical Society. It is expanded, but not very much. The mental apparatus in general, constitutional factors, clinical criteria, clssification, differential diagnosis and treatment form the subject-matter. So much in so short a space tends inevitably perhaps to generalizations without their supporting evidence and these are certainly not lacking. They are perhaps particularly unfortunate in a book that starts off with a very firm censure of the child analysts—Mrs Klein and her followers especially—on just these grounds of failure to produce evidence for their views. A great body of work by child analysts, which many feel to be a major contribution of psycho-analysis to child psychiatry, is in fact set aside.

To those already familiar with the author's other writings, in which his views are more fully supported, this book will be of interest. To others the succession of rather sweeping statements is likely to carry very much less conviction.

Fundamental Concepts in Clinical Psychology. By G. WILSON SHAFFER and RICHARD S. LAZARUS. New York. McGraw-Hill Book, Co. 1952. Pp. 540. $6.00.

This book sets out to cover the basic principles, methodology and general techniques that are the concern of the clinical psychologist, with a view to providing an introduction and supplement to more specialized studies. The field of the clinical psychologists, as here described, is clearly very much broader than would be normal in this country and indeed it is sometimes difficult to see what, if any, are its limits. A wide range of methods of psychological diagnosis and of therapy is discussed. To those entering on the study of mental health this book may prove useful, provided it is realized that the techniques in question are in fact the tools of a number of specialists rather than those of any one profession. There is a serviceable bibliography of the American literature with only scanty reference to that of other countries.

Psychiatric Aide Education. By BERNARD HALL *et al.* Menninger Clinic Monograph Series, No. 9. New York: Grune and Stratton. 1952. Pp. xviii + 168. No price given.

It appears that the 750,000 hospitalized mental patients in the U.S.A. are nursed by some 90,000 personnel of which only 10,000 are trained nurses. The rest are 'attendants', or 'psychiatric aides' as they are now called. Training schemes for psychiatric aides have existed in many American mental hospitals since about 1900, but these have had the effect of turning 'aides' into 'nurses' who, in turn, have tended to leave for more lucrative work elsewhere. In any case, only a small proportion of aides took the courses provided.

In order to help to remedy this state of affairs, the Menninger Aide School was opened in October 1949, which trained a total of sixty psychiatric aides in four intensive courses until it closed in 1952. The purpose of this School was to investigate the material that should be taught and the most suitable methods of teaching it, and also to determine the best selection and recruitment procedures. It was also hoped to graduate a small number of competent aides who would emerge as leaders for the psychiatric aide group. The book is a report of this investigation.

The results reported have only slight relevance to conditions in this country. In British mental hospitals the male 'attendant' is now unknown. All male nursing staff are either student or graduate mental nurses, and senior nursing staff usually have a general training as well. Female 'nursing assistants' are employed but, they always work under the direct supervision of fully trained senior nurses. Some mental hospitals have a training course for such nursing assistants, followed by an examination and a Hospital Certificate. The syllabus outlined in this book seems too advanced for nursing assistants yet not adequate for fully trained mental nurses.

RALPH HETHERINGTON

Diagnostic Casework in the Thomastic Pattern. By MARY J. MCCORMICK. New York: Columbia University Press. London: Geoffrey Cumberlege. 1954. Pp. xvi + 239. 28*s*. net.

It may be a matter of surprise to many psychologists that a theologian can contribute to our understanding and practice of case-work. But Miss McCormick proves otherwise. In this country we tend not to consider diagnostic case-work as a subject separate from either social case work or psychotherapy, but it is clear, if we make the distinction, that the social worker meets many problems which are neither strictly psychiatric nor administrative, but educational and moral. It is in these spheres that the profound human understanding of Thomist philosophy makes a great contribution, whether or not one follows it to its ultimate Christian conclusions. It is a philosophy which emphasizes at once the failings and the grandeur of human personality, and stresses the duty we have to each other to release the nobility from the pains of the failings, for 'the sum total of the Christian religion consists in mercy, as regards external works'. Miss McCormick has breadth and wisdom in her approach, while her counsel of prudence could well be more widely emulated.

Altogether, this is an excellent, if not a great book; and quite apart from its metaphysics, which will be stimulating to some and meaningless to others, it should prove of real value to the practical social worker. ADAM CURLE

Psychology, the Nurse and the Patient. By DORIS M. ODLUM. London: 'Nursing Mirror'. 1952. Pp. 114. 7*s.* 6*d.* net. 2nd edition, 1954. Pp. 168. 12*s.* 6*d.* net.

Dr Odlum's book is an attempt to fill the need for a comprehensive and readable introduction to psychology for nurses in training, and in this second edition, much new material has been included. Selection of topics and the amount of space and importance given to them are always difficult in a book of this kind, and the difficulty is particularly noticeable here. One could wish that more space had been given to purely psychological matters and less to details of nursing care, and that the new material had been incorporated better into the body of the book, so that repetitions and bad references were avoided. The most valuable part of the book is that on the nurse's relationships in her work, and for that reason, the book should prove useful for the nurse in training, but not as her only reading in psychology. A.M.S.

Issues in the Study of Talent. By D. W. BRAY. Columbia University Press; London: Geoffrey Cumberledge. 1954. Pp. xi+65. 16*s.*

It is difficult to understand why this small book was produced—particularly at such a price. It is intended to contribute towards the 'conservation of human resources' research project of the Columbia University Graduate School of Business, and sets out to 'develop a theoretical framework that will provide guide posts for future work'. It tries to do this by giving a brief and superficial review of the nature of intelligence, the work done on heredity and environment, the evidence of factor analysis and the relevance of motivation in work performance; then a 'research plan' is produced which depends very little upon the previous outline. The emphasis in the book is on the importance of environment: 'The average innate capacity of children in one socio-economic-cultural group does not differ substantially from the capacity of children in another group. Therefore, the wide differences in I.Q. among children from the different groups must be due primarily to differences in cultural opportunity'. A review of factor analysis leads to the conclusion that g reflects the cultural heritage of the society, and that Vernon's $v{:}ed$ and $k{:}m$ group factors arise from children using their 'general learning potential' either inside or outside school.

It is proposed 'to study individuals who have achieved prominence through their superior performance in music, physical science, or business administration'. The possibility of using professional baseball players is discussed and then—reluctantly?—abandoned. It is hoped to discover factors connected with the recognition, development and utilization of potential.

If the fashion of writing books describing what researches are *going* to be done becomes common, the task of reviewers is likely to become less and less attractive. STEPHEN WISEMAN

OTHER PUBLICATIONS RECEIVED

Emil A. Gutheil. *The Handbook of Dream Analysis*. New York: Liveright Publishing Corporation; London: George Allen and Unwin. 1952. Pp. 710. 48*s*. net.

Interrelations of Cultures. Paris: U.N.E.S.C.O. 1953. Pp. 387. 14*s*. 6*d*. net.

Kurt Reidemeister. *Die Unsachlichkeit der Existenzphilosophie*. Berlin: Springer-Verlag. 1954. Pp. iv+40. DM 4.80.

Kurt Reidemeister: *Geist und Wirklichkeit. Kritische Essays*. Berlin: Springer-Verlag. 1953. Pp. iv+92. DM. 8.60.

Günter Clauser. *Die Kopfuhr; Das automatische Erwachen*. Stuttgart: Ferdinand Enke Verlag. 1954. Pp. viii+108. DM 11.80.

Peter R. Hofstätter. *Einführung in die Sozialpsychologie*. Stuttgart: Humboldt-Verlag. 1954. Pp. 536. DM. 12·50.

Index Bibliographicus. Vol. ii. compiled by Theodore Besterman. Paris: U.N.E.S.C.O. 1952. (3rd edition.) Pp. xi+72. 9*s*. 6*d*. net. (H.M.S.O.)

Ernest, L. Remits. *The Feeling of Superiority and Anxiety—Superior*. Ottawa: Runge Press Ltd. 1953. Pp. x+67.

Pierre Pichot. *Les Tests Mentaux*. Paris: Presses Universitaires de France. 1954. Pp. 128.

Aids to Educational Research. (New Edition.) Publications of the Scottish Council for Research in Education xxxi. London: University of London Press. 1954. Pp. viii +42. 2*s*. 6*d*. net.

The Teaching of the Social Sciences in the United States. Paris: U.N.E.S.C.O. 1954. Pp. 150. 6*s*. net.

William V. Silverberg, M.D. *Childhood Experience and Personal Destiny*. New York: Springer Publishing Company, Inc. 1952. Pp. xiv+289. $4.50.

Joseph Nuttin. *Psychoanalysis and Personality*. London: Sheed and Ward. 1954. Pp. xiv+310. 16*s*. net. (This is an English translation of the book reviewed in the *British Journal of Psychology*, vol. xliii, pt. 3, August 1952.)

A General Selection from the Works of Sigmund Freud. Edited by John Rickman (Psycho-Analytical Epitomes No. 1.) London: The Hogarth Press. 1953. Pp. x+332. 10*s*. 6*d*. net.

Sigmund Freud. *Civilisation, War, and Death*. Edited by John Rickman. (Psycho-Analytical Epitomes No. 4). London: The Hogarth Press. 1953. Pp. xviii+118. 8*s*. 6*d*. net.

Cultural Patterns and Technical Change. Edited by M. Mead. Paris: U.N.E.S.C.O. 1953. Pp. ii+350. 10*s*. 6*d*. net.

✓

1962